



Imagerie à travers la turbulence : mesure inverse du front d'onde et centrage optimal

Xavier Rondeau

► To cite this version:

Xavier Rondeau. Imagerie à travers la turbulence : mesure inverse du front d'onde et centrage optimal. Modélisation et simulation. Université Claude Bernard - Lyon I, 2007. Français. NNT : . tel-00220467

HAL Id: tel-00220467

<https://theses.hal.science/tel-00220467>

Submitted on 28 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre: 240 - 2007

Thèse de Doctorat en Physique et Astrophysique

présentée

devant l'Université Lyon 1 - Claude Bernard

par Xavier RONDEAU

pour obtenir le diplôme de doctorat

École Doctorale : Physique et Astrophysique (Phast ENS-Lyon)

Équipe d'accueil : AIRI (CRAL-Observatoire de Lyon)

Imagerie à travers la turbulence : mesure inverse du front d'onde et centrage optimal

Soutenue le 12 Novembre 2007 devant la commission d'examen composée de

MM.:	Olivier	MICHEL	Rapporteur
	Laurent	MUGNIER	Rapporteur
MM.:	Nicholas	DEVANEY	Examineur
	Philippe	PRUGNIEL	Examineur
	Gérard	ROUSSET	Examineur
MM.:	Renaud	FOY	Directeur de thèse
	Eric	THIEBAUT	Directeur de thèse

Remerciements

Bien qu'une thèse soit censée dépendre en grande partie de celui qui s'y attelle, en pratique de nombreux autres facteurs sont tout aussi décisifs : la pertinence et l'originalité du sujet de thèse, l'interaction avec les encadrants, le soutien des amis et de la famille et cætera. Cela étant dit certaines personnes sont aussi capables de se frayer leur chemin à travers n'importe quelle tempête ... Malheureusement ou bienheureusement pour moi ? Je ne peux pas me revendiquer pour ma thèse de cette dernière catégorie, car j'ai bénéficié de conditions très favorables à son bon déroulement.

Je souhaite commencer par remercier ma chérie pour sa présence à mes côtés tout au long de ces trois années de thèse. Je pense qu'elle a bien du mérite d'avoir su avec bienveillance digérer mes sautes d'humeur tout en tentant parfois de les apaiser, d'autant que ses études ne sont pas de tout repos non plus ! Mes parents et ma soeur qui ont toujours eu l'intelligence de ne pas se décharger de leurs problèmes sur moi, et de me prêter l'oreille attentive, compréhensive et apaisante associée à leur plus grande expérience. Sans parler de l'aide matérielle qui m'a permis de vivre cette expérience avec plus de confort qu'il n'est normalement possible.

J'ai également eu beaucoup de chance concernant mon équipe de recherche. Une équipe protéiforme : créative et pluridisciplinaire tout en étant à la pointe de ses réalisations ; aux individualités fortes et tranchées mais malgré tout complémentaires et soudées. Un environnement très riche donc, pour effectuer ma thèse. Je veux avant tout remercier Eric (Thiébaud) qui a le plus contribué à cette thèse et à mon apprentissage, nonobstant le grand-écart à effectuer par rapport à ses compétences. Jusques alors peu convaincu par l'intérêt de mes études scientifiques, je lui dois en effet d'avoir pu changer d'avis au contact de sa joie communicative dans l'analyse et la modélisation scientifique. Je reste admiratif de sa capacité à communiquer ce qu'il pense en associant à ses grandes facultés d'analyse et à sa vivacité d'esprit, une subtile pédagogie qui ne sacrifie rien à la rigueur de son propos. Côtayer ce très rare spécimen aura donc été une véritable aubaine. Je tiens ensuite à remercier Renaud (Foy) pour m'avoir accueilli dès mon stage de master dans l'équipe Airi sur des sujets aussi ambitieux que pluridisciplinaires et innovants. Il a le grand mérite de s'aventurer avec une persévérance édifiante sur des chemins que d'aucuns jugeraient trop pentus. Je souhaite au passage au projet ELP-OA qu'il a su défendre face à de nombreux écueils, les francs succès qu'il mérite lors de la phase expérimentale qui vient de débiter. Je remercie également Michel (Tallon) des discussions que j'ai pu avoir avec lui malgré sa grande charge de travail, ainsi que de son aide méthodologique. Sa rigueur dans l'organisation et l'agencement a de quoi donner des complexes. Notamment sa démarche et sa capacité à retourner aux fondements de chaque problème pour ensuite agencer ses briques les unes après les autres, sont évidentes et exemplaires.

Bien que je n'ai sans nul doute pas été assez serein et actif dans les relations du quotidien au sein du laboratoire, j'ai pu cependant ressentir la bienveillance du reste

de l'équipe (Isabelle, Clémentine, Ferréol) que je remercie également et notamment Françoise (Foy) pour son souci des autres et ses efforts pas toujours récompensés pour renforcer la convivialité au sein de l'observatoire.

Je ne saurais oublier de remercier Johan (Kosmalski) dont la complicité au sein du laboratoire aura été d'un grand réconfort pour ma thèse, lorsque la marmite menaçait d'exploser par exemple. Julien (Girard), joyeux luron avec lequel j'ai partagé mon bureau lors de ma première année dans l'équipe et qui a bien facilité mon entrée en la matière. Je pourrais m'épancher beaucoup plus, mais comme ce n'est je pense pas le bon endroit pour cela, je souhaite pour finir remercier brièvement et symboliquement tous les amis qui en dehors de l'observatoire m'ont apporté le réconfort nécessaire au quotidien (Delphine, Boris, Etienne, la famille de ma Flo, ...).

D'avoir évolué dans un environnement si favorable ne me permet donc pas de concrétiser ce fantasme de remercier les autres pour avoir donné à mon intelligence l'occasion de briller par leur opposition ; mais bien au contraire me sert d'exemple à suivre pour m'améliorer. Sincèrement donc :

m e r c i !

Table des matières

Table des matières	-1
Introduction	5
1 Haute résolution angulaire pour l'imagerie à travers la turbulence	9
1.1 Imagerie astronomique basée sur la Terre	10
1.1.1 Relation fondamentale entre objet et image	10
1.1.2 Images au foyer du télescope	11
1.1.2.1 Onde incidente au télescope	12
1.1.2.2 Image formée par le télescope	12
1.2 Turbulence Atmosphérique	14
1.2.1 Description physique de la turbulence	14
1.2.2 Description statistique	15
1.2.2.1 Hypothèses statistiques	15
1.2.2.2 Fluctuations spatiales de l'indice de réfraction d'une couche turbulente	15
1.2.2.3 Fluctuations spatiales de la phase sur la pupille . . .	16
1.2.2.4 Fluctuations temporelles de la phase sur la pupille .	17
1.3 Effets de la turbulence dans les images	18
1.3.1 Effets dans les images longues poses	18
1.3.1.1 Fonction de transfert longue pose	18
1.3.1.2 Paramètre de Fried	18
1.3.1.3 Résolution angulaire longue pose	19
1.3.2 Effets dans les images courtes poses	19
1.3.2.1 Images tavelées	19
1.3.2.2 Cohérence temporelle	20
1.3.2.3 Fonction de transfert courte pose	20
1.3.3 Illustration	21
1.4 Techniques pour obtenir les Hautes Résolutions Angulaires dans les images longues poses	22
1.4.1 Techniques de HRA a posteriori	22

1.4.2	Optique Adaptative	24
1.4.2.1	Corrections et estimations du front d'onde en temps réel	25
1.4.3	Source de référence pour les techniques HRA	27
1.4.3.1	Angle d'isoplanétisme	27
1.4.3.2	Etoile laser artificielle	29
1.4.4	Etoile laser polychromatique	30
1.4.4.1	Dépendance chromatique du chemin optique	30
1.4.4.2	Mesure chromatique différentielle de la pente de la surface d'onde	31
1.4.4.3	Cascade radiative de l'étoile laser polychromatique	33
2	Introduction aux problèmes inverses	37
2.1	Préambule	37
2.1.1	Description phénoménologique	37
2.1.2	Les débuts historiques de formalisation	38
2.2	Formalisation probabiliste	38
2.2.1	Induction	39
2.2.2	Espaces probabilisés des observables et des paramètres	39
2.2.3	Densité de probabilité expérimentale	40
2.2.4	Densité de probabilité a priori	40
2.2.5	Densité de probabilité a posteriori	40
2.3	Remarques	41
2.3.1	Choix du maximum	41
2.3.2	Importance du volume choisi	41
2.3.3	Fonction d'adéquation a posteriori	42
2.3.4	Poids de la régularisation	42
2.4	Cas particulier des fluctuations gaussiennes	42
2.5	Optimisation pseudo-continue pour les problèmes inverses non-linéaires	43
2.5.1	Formalisme	43
2.5.2	Algorithme de Newton	44
2.5.3	Algorithmes de quasi-Newton	45
2.5.3.1	Algorithme de Levenberg-Marquardt	46
2.5.3.2	Algorithme de quasi-Newton partitionné	46
2.5.3.3	Algorithmes de métrique variable.	47
2.5.4	Algorithme des gradients conjugués	48
2.5.5	Ajustement de la longueur de pas	49
2.5.5.1	Conditions de Wolfe	49
2.5.5.2	Région de confiance	50
2.6	Exploration discontinue de type Monte-Carlo	52
2.6.1	Débuts historiques	52

2.6.2	Application aux problèmes inverses	52
2.6.3	Chaînes de Markov	53
2.6.3.1	Algorithme de Métropolis	54
2.6.4	Recuit simulé	55
3	Reconstruction de phase à partir d'une image tavelée	57
3.1	Historique	57
3.2	Adéquation expérimentale (ML)	58
3.2.1	Modèle direct des images	58
3.2.2	Bruits et incertitudes expérimentales	59
3.2.3	Adéquation expérimentale	60
3.3	Adéquations a priori	60
3.3.1	Ajustement analytique non-ambigu du paramètre d'amplitude	60
3.3.2	Ambiguïtés de la phase dans l'adéquation expérimentale . . .	61
3.3.2.1	Ambiguïté stricte du signe de la composante paire de la phase	61
3.3.2.2	Ambiguïté de signe de chaque mode pair de la phase	62
3.3.2.3	Ambiguïtés strictes de périodicité modulo 2π de la phase	62
3.3.3	Adéquation a priori des paramètres de phase dépistonnée et détiltée	63
3.3.3.1	Densité de probabilité a priori	63
3.3.3.2	A priori alternatif de lissage	64
3.3.4	Ambiguïtés de tip-tilt	65
3.3.5	Régularisation du tip-tilt	65
3.4	Reparamétrisation	66
3.4.1	Importance de la reparamétrisation pour le problème inverse .	66
3.4.2	Reparamétrisation modale de la phase dépistonnée et détiltée .	66
3.4.3	Reparamétrisation zonale de la phase dépistonnée et détiltée .	68
3.4.4	Restriction du nombre de paramètres	68
3.5	Algorithme d'optimisation locale	69
3.5.1	Choix du type d'algorithme d'optimisation non linéaire	69
3.5.2	Ajustement automatique et à la volée du poids de la régularisation	69
3.5.3	Projection dans un sous-espace de directions privilégiées . . .	70
3.5.4	Région de confiance	72
3.5.5	Calcul rapide des dérivées partielles	73
3.6	Simulation de données	73
3.7	Optimisation globale de phases décorréliées	75
3.7.1	Stratégie pseudo-aléatoire	75
3.7.2	Corrélation entre f_{exp} et l'erreur de reconstruction de la phase	76
3.7.3	Efficacité de la stratégie en fonction du D/r_0	78

3.7.4	Efficacité de la stratégie en fonction du bruit	78
3.7.5	Convergence finale de l'algorithme	79
3.8	Optimisation globale pour des phases à forte corrélation temporelle . .	80
3.9	Qualification de l'importance de certains choix	82
3.9.1	Bornage du tip-tilt	82
3.9.2	Régularisation de la phase	82
3.9.3	Echantillonnage de la phase	84
3.9.4	Région de confiance	84
3.10	Conclusions	86
4	Diversité de phase polychromatique	89
4.1	Modélisation polychromatique	90
4.1.1	Modèle des images	90
4.1.2	Critère d'adéquation a posteriori polychromatique	90
4.1.3	Algorithme d'optimisation	91
4.2	Bras de levier polychromatique	91
4.3	Ajustement des poids relatifs des critères monochromatiques	94
4.4	Conclusion	95
5	Centrage	97
5.1	Estimation de la position d'une source ponctuelle à partir d'images tavelées	97
5.1.1	Différence entre barycentre et pente de la surface d'onde . . .	97
5.1.2	Erreurs et biais affectant le centrage	98
5.1.2.1	Bruit de discrétisation spatiale	99
5.1.2.2	Bruit du détecteur	100
5.1.2.3	Bruit de photons	100
5.1.2.4	Troncature par le détecteur	101
5.1.3	Borne de Cramér-Rao	102
5.1.4	Estimation optimale du déplacement à partir des images tavelées	103
5.1.4.1	Borne de Cramér-Rao pour l'estimation du centre des images tavelées sans connaissance de la phase .	103
5.1.4.2	Bénéfice attendu de la connaissance de la phase . .	104
5.1.5	Simulations	104
5.1.5.1	Barycentre	105
5.1.5.2	Centrage à phase connue	107
5.1.5.3	Gain de précision pour le centrage	110
5.2	Estimation de la pente différentielle par interspectre bichromatique . .	112
5.2.1	Principe de la mesure	112
5.2.2	Simulations	114

5.3	Amélioration de la résolution spatiale d'un détecteur à comptage de photons	117
5.3.1	Présentation des caméras CPng/Algol	117
5.3.2	Evènements parasites	118
5.3.3	Modèle de maximum de vraisemblance pour la détection des photoévènements	121
5.3.4	Détection temps réel	123
5.3.5	Recentrage sous-pixel des photoévènements	125
5.3.6	Image des microcanaux	126
Bibliographie		137

Introduction

Démesurément hors de portée des sens humains, l'univers permet de faire évoluer notre rapport au réel et à l'imaginaire. La révolution entamée au début du XVII^{ème} siècle par Galilée dans l'observation des objets astronomiques à l'aide d'un instrument optique plus puissant que l'oeil, voit aujourd'hui poindre la construction d'une nouvelle génération de télescopes monolithiques de très grand diamètre $D \geq 30$ m, avec le TMT (Thirty Meter Telescope) des Etats-Unis et l'E-ELT (European Extremely Large Telescope) européen. Grâce à leur plus grand diamètre, ces télescopes permettront à la fois d'étudier des objets beaucoup moins brillants, et également d'en percevoir beaucoup plus de détail, puisque leur résolution angulaire $\lambda/D \approx 3$ millisecondes d'arc à 500 nm est typiquement 5 fois supérieure à celle de la génération actuelle de télescopes. Malheureusement les interactions fluctuantes de la lumière avec le milieu turbulent atmosphérique traversé, dégradent typiquement cette capacité de détail à celle d'un télescope de diamètre équivalent $r_0 \approx 15$ cm pour $\lambda = 500$ nm, et cela pour les meilleurs sites d'observation astronomique sur Terre – Mauna Kea à Hawaï ou le Dôme C en Antarctique par exemple.

La première solution pour s'affranchir des dégradations de la résolution angulaire liées à l'atmosphère consiste à les éviter en satellisant des télescopes spatiaux, comme Hubble et son successeur le JWST. Cela permet également d'éviter les aléas climatiques ; de placer les télescopes dans un environnement thermodynamique plus stable qu'au sol ; et d'éviter le filtrage du spectre électromagnétique (EM) par l'atmosphère, lequel affecte fortement les rayonnements à haute énergie à partir des ultra-violets – bienheureusement pour la survie animale. Toutefois ces télescopes spatiaux coûtent au moins dix fois plus cher que les télescopes terrestres, et il est beaucoup plus difficile d'assurer leur maintenance ou de les équiper au fur et à mesure avec l'instrumentation propre à chaque mission scientifique. Il semble donc qu'il faille attendre un nouvel élan dans la conquête spatiale, une fois les objectifs de celle-ci mieux délimités, et cela en dépit même du franc succès du télescope spatial Hubble, pour voir fleurir les télescopes spatiaux en plus grand nombre.

La deuxième solution pour retrouver toute la capacité potentielle de détail des télescopes de grand diamètre, consiste à corriger au sol les dégradations fluctuantes causées par l'atmosphère sur la phase de l'onde EM principalement. De nombreuses techniques, réunies sous le nom de techniques de Haute Résolution Angulaire (HRA),

ont été proposées dans ce but. Elles peuvent globalement être différenciées en deux grandes familles, selon que les mesures et corrections sont appliquées physiquement sur la pupille du télescope à l'aide d'une Optique Adaptative fonctionnant plus rapidement que le temps caractéristique d'évolution de la turbulence ; ou bien alors que les corrections sont modélisées et utilisées mathématiquement pour extraire directement l'information HRA d'une série d'images courtes poses, par une formulation en *problème inverse* comme la déconvolution myope/aveugle par exemple. Dans tous les cas, le flux de référence servant pour l'analyse des dégradations induites par la turbulence doit être relativement important, ce qui limite en pratique le domaine angulaire du ciel pour lequel des corrections peuvent être effectuées à moins de 1% dans le visible.

Pour remédier à ce problème de couverture insuffisante du ciel en source de référence suffisamment brillante, Foy and Labeyrie [1985] ont proposé le concept d'étoile laser artificielle de référence, qui est aujourd'hui largement accepté et utilisé. Malheureusement les portions de couches turbulentes traversées par la lumière de l'étoile artificielle et par celle de l'objet observé à l'infini ne sont pas les mêmes. La tomographie de la turbulence à l'aide de plusieurs sources artificielles de référence a ainsi été proposée et s'est imposée pour la Haute Résolution Angulaire en astronomie en parallèle de l'étoile laser. La tomographie permet de surcroît d'élargir le champ du télescope bénéficiant des corrections. Cependant il subsiste encore le problème majeur de l'indétermination de la position de l'étoile laser dans le ciel. En effet on ne sait pas séparer les composantes de déplacement aller et retour de la lumière de l'étoile laser artificielle causées par la turbulence. Dans le cas où le laser est tiré par le télescope d'observation on voit même que ces deux composantes s'annulent d'après le principe de retour inverse de la lumière. Cette indétermination est concrètement dramatique, puisque plus de 80% de la variance de la phase qu'on cherche à compenser pour obtenir la HRA des télescopes de grand diamètre est liée à ce seul bouger de l'étoile. Le concept d'Etoile Laser Polychromatique (ELP) a ainsi été proposé par Foy [1995], et consiste à estimer ce déplacement par une mesure différentielle chromatique. Le projet ELP-OA, qui débute sa phase de démonstration expérimentale, s'appuie ainsi sur la dépendance chromatique de l'indice de réfraction de l'air — qui est également valable pour le déplacement de l'étoile causé par l'atmosphère — et l'utilisation de flux lumineux à plusieurs longueurs d'onde en provenance d'une étoile laser artificielle polychromatique.

Dans un premier chapitre, je rappelle les notions fondamentales de l'imagerie à travers la turbulence qui me servent tout au long de ce mémoire de thèse. J'insiste en particulier sur les effets statistiques de la turbulence sur la phase de l'onde électromagnétique qui la traverse et sur les conséquences pour la transmission des fréquences spatiales des objets dans les images longues poses ou courtes poses.

Je présente ensuite une introduction aux problèmes inverses avec un formalisme et des notions qui me servent ensuite pour introduire la formulation statistique robuste associée au problème de reconstruction de phase. Je dresse également un aperçu des

différentes techniques d'optimisation non-linéaire afin de justifier les choix particuliers et les développements effectués par la suite pour l'optimisation globale du critère non-linéaire et très multi-modal d'inférence qui est issu de cette formulation.

Dans le chapitre trois, je détaille le modèle et l'algorithme d'optimisation globale que j'ai développés pour le problème inverse non-linéaire de reconstruction de phase à partir d'une seule image tavelée. Je caractérise mon algorithme pour diverses conditions de turbulence et de bruit.

Dans le chapitre suivant, j'introduis la technique de diversité polychromatique de phase pour pouvoir gérer plus de degrés de liberté, c'est à dire des conditions du turbulence plus difficiles ou bien un diamètre de télescope plus grand pour un r_0 constant. Je développe en particulier un moyen automatique et approprié d'ajuster le poids relatif des différentes longueurs d'onde au cours de l'optimisation, afin d'assurer l'optimisation globale du critère de maximum de vraisemblance pénalisé qui est ici fortement multi-modal.

Enfin dans un dernier temps, je développe des méthodes pour améliorer le centrage d'une tache image. Je montre que la modélisation de la distribution de la tache image apportée par l'estimation de toutes les aberrations de phase permet de gagner en précision de centrage par rapport au barycentre, ce qui s'applique en particulier ici à l'estimation de la position de l'étoile laser à partir d'images tavelées. Pour le cas où l'étoile laser polychromatique créée serait très étendue et résolue par le télescope d'observation, je développe également une méthode d'estimation du déplacement chromatique différentiel, basée sur l'interspectre bichromatique des images tavelées et visant à exploiter un nombre de fréquences spatiales intermédiaire entre la reconstruction de phase et le barycentre. Enfin je présente un algorithme inverse pour traiter en temps réel et de façon optimale les données d'un détecteur à comptage de photons et je montre ses bénéfices en terme de résolution spatiale à partir de vraies données.

Chapitre 1

Haute résolution angulaire pour l'imagerie à travers la turbulence

L'analyse d'un objet à partir de la détection de l'onde électromagnétique (EM) que cet objet émet ou réémet est un ingrédient essentiel de la prise d'information humaine. L'emploi de ce procédé s'est donc instinctivement développé pour caractériser des objets à distance et des milieux de propagation, et en particulier pour les observations astronomiques, pour lesquelles l'oeil humain ne suffit largement pas. Aujourd'hui la caractérisation des objets astronomiques à partir de la détection d'autres signaux physiques comme les ondes gravitationnelles, que l'homme appréhende plus difficilement, commence tout juste à se développer. Je ne m'intéresse toutefois ici qu'à la caractérisation du milieu turbulent que représente l'atmosphère terrestre pour les images astronomiques au sol. Ces images sont la répartition de l'énergie de l'onde EM qui a traversé l'atmosphère, dans un plan focal de détection $\mathbf{x}=(x, y)$. Elles dépendent également de la longueur d'onde λ et du temps t .

Dans ce premier chapitre, je décris brièvement le contexte de la Haute Résolution Angulaire (HRA) pour l'imagerie astronomique terrestre longue pose, dans lequel s'inscrivent mes travaux de thèse sur le centrage et sur la reconstruction de phase.

Je commence dans une première section par introduire les relations générales de l'imagerie, et je détaille l'expression analytique des images observées au foyer du télescope en fonction de l'onde électromagnétique sur la pupille d'entrée du télescope, car cette modélisation me sert en particulier pour la reconstruction de phase.

Je rappelle ensuite les propriétés générales de la turbulence, et ses effets sur l'observation astronomique. Je réintroduis pour cela la statistique de Kolmogorov sur les fluctuations spatio-temporelles de l'onde électromagnétique qui traverse la turbulence. J'utilise cette statistique dans toute la suite de ma thèse mais les méthodes que je développe n'en dépendent pas. Celles-ci peuvent donc tout aussi bien être utilisées avec des statistiques décrivant la turbulence plus finement ou autrement si nécessaire. Je discute aussi dans cette section de la perte ou de l'atténuation des détails, causé par la

turbulence dans les images longues poses ou courtes poses.

Je rappelle finalement brièvement dans la dernière section les principales techniques HRA qui sont utilisées pour récupérer le détail potentiel des télescopes de grand diamètre, et en particulier le projet ELP-OA d'Etoile Laser Polychromatique pour l'Optique Adaptative. Ces techniques sont décrites plus en détail dans de nombreux ouvrages de référence, et notamment [Roggemann et al., 1997; Roddier, 1999; Foy and Foy, 2005].

1.1 Imagerie astronomique basée sur la Terre

1.1.1 Relation fondamentale entre objet et image

Soit $i_{\lambda,t}(\mathbf{x})$ la distribution instantanée et monochromatique d'énergie détectée dans un plan \mathbf{x} , et $o_{\lambda,t}(\alpha)$ la distribution d'énergie associée de l'objet, observé dans une direction angulaire α . Soit $h_{\lambda,t}$ le modèle d'imagerie qui relie entre elles ces deux distributions, et qui prend en compte dans notre cas les effets de la propagation dans un milieu turbulent et les effets de l'instrument d'imagerie, et ne contient pas dispersion spectrale ou temporelle. En considérant que l'objet émet une lumière spatialement incohérente, l'image s'écrit comme la somme des images de chacun des points de l'objet :

$$i_{\lambda,t}(\mathbf{x}) = \int h_{\lambda,t}(\mathbf{x}, \alpha) \times o_{\lambda,t}(\alpha) d\alpha. \quad (1.1)$$

Cette relation se simplifie sous la forme d'un produit de convolution, dans la zone angulaire où le modèle peut être considéré comme invariant par translation (domaine d'isoplanétisme) :

$$i_{\lambda,t}(\mathbf{x}) = \int h_{\lambda,t}(\mathbf{x} - \alpha) \times o_{\lambda,t}(\alpha) d\alpha = [h_{\lambda,t} \star o_{\lambda,t}](\mathbf{x}), \quad (1.2)$$

et s'écrit encore plus simplement dans l'espace de Fourier associé sous la forme :

$$I(\mathbf{u}) = H(\mathbf{u}) \times O(\mathbf{u}), \quad (1.3)$$

où \mathbf{u} est la fréquence spatiale conjuguée à \mathbf{x} , et où on a utilisé les lettres capitales pour désigner les transformées de Fourier. Cette dernière relation de filtrage exprime l'atténuation de l'information de l'objet dans l'image par la fonction de transfert H (MTF), en fonction de la fréquence spatiale \mathbf{u} . En pratique on ajoute souvent un terme additif $n(\mathbf{x})$ dans l'expression (1.2) pour modéliser les fluctuations physiques de la relation, si bien que l'équation (1.3) se réécrit $I(\mathbf{u}) = H(\mathbf{u}) \times O(\mathbf{u}) + N(\mathbf{u})$. La transmission de l'information de l'objet diminue le plus souvent avec l'augmentation de la fréquence spatiale, jusqu'à une fréquence de coupure à partir de laquelle le système d'imagerie ne

transmet plus les fréquences spatiales de l'objet, ou bien de façon négligeable devant les contributions $N(\mathbf{u})$ du bruit.

Dans le cas des images astronomiques longues poses, l'atmosphère et le télescope sont modélisés comme deux filtres spatiaux cascades, pourvu que le champ d'observation ne dépasse pas le domaine d'isoplanétisme. Autrement dit la fonction de transfert longue pose $\langle H(\mathbf{u}) \rangle$ s'exprime comme le produit de la fonction de transfert longue pose de l'atmosphère $B(\mathbf{u})$ et de la fonction de transfert du télescope $T(\mathbf{u})$ (supposée invariante) :

$$\langle H(\mathbf{u}) \rangle = B(\mathbf{u}) T(\mathbf{u}) . \quad (1.4)$$

La fonction de transfert du télescope $T(\mathbf{u})$ est l'autocorrélation normalisée de l'amplitude complexe sur la pupille du télescope. En effet ce résultat est obtenu d'après le théorème de Wiener-Khintchine parce que l'image au foyer du télescope est la densité spectrale de l'amplitude complexe sur la pupille (ce qui est redémontré dans la section suivante). En l'absence de turbulence et dans le cas idéal où le télescope n'a pas de défaut optique, l'amplitude complexe sur la pupille est une fonction réelle $P(\mathbf{u})$ appelée fonction pupille, qui vaut 1 sur la partie réfléchissante de la pupille et 0 en dehors. La fonction de transfert du télescope s'écrit donc :

$$T(\mathbf{u}) = \frac{\iint P(\mathbf{u}') P(\mathbf{u}' + \mathbf{u}) d\mathbf{u}'}{\iint |P(\mathbf{u}')|^2 d\mathbf{u}'} \quad (1.5)$$

On voit que la fonction de transfert du télescope s'annule quelle que soit la forme de la pupille, à partir de la fréquence de coupure $u_c = D/\lambda$, où D est la dimension maximale de la pupille. La séparation angulaire maximum du détail des objets ou *résolution angulaire* du télescope s'écrit donc $\alpha_c = 1/u_c = \lambda/D$ et s'améliore proportionnellement avec l'augmentation du diamètre D du télescope.

Comme je le montre plus en détail à la section 1.3 après avoir introduit les effets statistiques de la turbulence, le filtre $B(\mathbf{u})$ que constitue l'atmosphère coupe tout le détail potentiel des télescopes de grand diamètre dans les images longues poses, si les effets délétères de la turbulence ne sont pas corrigés.

1.1.2 Images au foyer du télescope

L'instrument aujourd'hui le plus utilisé pour former des images astronomiques est le télescope. Comme la lumière astronomique est particulièrement raréfiée, la surface collectrice de diamètre D du télescope est tout d'abord maximisée pour obtenir un rapport signal à bruit suffisamment élevé. Elle l'est également, comme on vient de le voir, pour conserver le maximum de détail des objets observés. C'est pourquoi la construction actuelle de télescopes de grands diamètres avec des systèmes pour corriger les effets de la turbulence constitue une avancée majeure pour l'observation de notre univers.

La formation des images est traitée de façon détaillée dans [Goodman, 1968; Born and Wolf, 1999]. Je rappelle ici uniquement les grandes lignes de ces développements, pour aboutir à l'expression analytique ou *modèle direct* exprimant l'image au foyer du télescope en fonction de l'amplitude complexe de l'onde EM incidente sur le miroir primaire du télescope.

1.1.2.1 Onde incidente au télescope

Après propagation dans le milieu interstellaire quasi-vide, dans la direction \mathbf{z} et sur une distance astronomique, l'onde sphérique émise par un point source au temps t_0 , s'exprime d'après les équations de Maxwell dans un plan \mathbf{r} et à l'instant t sous la forme d'une onde plane électromagnétique :

$$A_\lambda(\mathbf{r}, t) = \alpha_\lambda \times e^{j(k_\lambda \mathbf{z} \cdot \mathbf{r} - \omega_\lambda t + \text{cste})}, \quad (1.6)$$

où $k_\lambda = 2\pi n_\lambda / \lambda = \omega_\lambda / c$ est le nombre d'onde, n_λ est l'indice de réfraction du milieu de propagation ($n_\lambda^{\text{vide}} = 1$) et ω_λ est la pulsation de l'onde. Cette expression ne dépend pas de la position dans le plan \mathbf{r} , lorsque celui-ci est perpendiculaire à la direction de propagation.

L'onde plane interagit ensuite de façon hétérogène avec le milieu atmosphérique terrestre qu'elle traverse, parce que la quantité de matière est distribuée de façon turbulente. L'onde électromagnétique au sol dans un plan \mathbf{r} de réception est donc affectée de fluctuations spatiales et temporelles et s'écrit :

$$A_\lambda(\mathbf{r}, t) = \alpha_\lambda(\mathbf{r}, t) e^{j(\varphi_\lambda(\mathbf{r}, t) - \omega_\lambda t + \text{cste})}, \quad (1.7)$$

où $\alpha_\lambda(\mathbf{r}, t)$ prend en compte les absorptions et diffusions fluctuantes du milieu atmosphérique et où le retard $\varphi_\lambda(\mathbf{r}, t)$ traduit le temps hétérogène de traversée de l'atmosphère. En faisant l'hypothèse que la turbulence atmosphérique est organisée en strates indépendantes et quasi-perpendiculaires à la propagation de l'onde EM, le retard au sol s'écrit

$$\begin{aligned} \varphi_\lambda(\mathbf{r}, t) &= \frac{2\pi}{\lambda} \times \int_{C(\mathbf{r})} n_{\lambda,t}(\mathbf{c}) d\mathbf{c} \\ &= \frac{2\pi}{\lambda} \times \delta_\lambda(\mathbf{r}, t), \end{aligned} \quad (1.8)$$

où le chemin optique $\delta_\lambda(\mathbf{r}, t)$ traduit les fluctuations de l'indice de réfraction le long du chemin le plus court $C(\mathbf{r})$ emprunté par la lumière.

1.1.2.2 Image formée par le télescope

En adoptant le point de vue d'Huygens-Fresnel de la propagation de proche en proche des ondes électromagnétiques, à l'endroit de la rencontre d'un obstacle, la diffraction est décrite par la somme des ondes sphériques émises par des sources secondaires au niveau de l'obstacle (intégrale de Rayleigh-Sommerfeld). Dans le cas du

télescope le premier obstacle rencontré est le miroir primaire. L'amplitude complexe diffractée par ce miroir à une distance z dans le plan décrit par \mathbf{x} , s'écrit :

$$a_{\lambda}^{\text{diff}}(z, \mathbf{x}, t) = \frac{1}{j\lambda} \iint_D P_{\lambda}(\mathbf{r}) A_{\lambda}(\mathbf{r}, t) \frac{e^{-jk_{\lambda}d}}{d} \cos(\mathbf{e}_z, \mathbf{e}_d) d\mathbf{r}, \quad (1.9)$$

où les vecteurs \mathbf{e}_z et \mathbf{e}_d représentent respectivement la normale au plan de l'obstacle

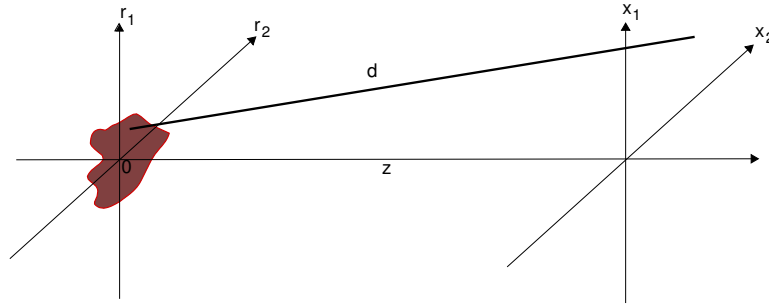


FIG. 1.1 – Amplitude diffractée

et la direction associée à la distance $d = \sqrt{|\mathbf{x} - \mathbf{r}|^2 + z^2}$, et où $P_{\lambda}(\mathbf{r})$ est la transmission d'amplitude de la pupille. Dans le cas du télescope on effectue les hypothèses simplificatrices de l'approximation paraxiale ($\cos(\mathbf{e}_z, \mathbf{e}_d) \simeq 1$) et de la diffraction à l'infini de Fraunhofer ($z \gg \frac{\pi|\mathbf{x}|^2}{\lambda}$). Avec $d \simeq z(1 + \frac{1}{2} \frac{|\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{r}}{z^2})$ on obtient l'expression de l'amplitude diffractée à l'infini par le miroir primaire du télescope :

$$a_{\lambda}^{\text{diff}}(\mathbf{x}, t) = \frac{\exp(-jk_{\lambda}z) \times \exp(-j\pi|\mathbf{x}|^2/\lambda z)}{j\lambda z} \times \iint_D P_{\lambda}(\mathbf{r}) A_{\lambda}(\mathbf{r}, t) e^{j2\pi\mathbf{x} \cdot \mathbf{r}/\lambda z} d\mathbf{r}. \quad (1.10)$$

En pratique l'éloignement entre les plans du miroir primaire et du foyer du télescope n'est pas grand par rapport à D , mais un miroir secondaire vient ramener l'amplitude diffractée à l'infini au foyer du télescope. La diffraction de Fresnel par le second miroir, se traduit simplement par la multiplication par une exponentielle complexe de transmission devant l'intégrale de l'équation (1.10). L'image au foyer du télescope dans le plan \mathbf{x} est l'énergie de l'onde EM, c'est à dire le module carré intégré autour du temps t de l'amplitude complexe à cet endroit. En reconnaissant la transformée de Fourier \mathcal{F} dans l'expression (1.10), et en notant $\mathbf{u} = \mathbf{r}/(\lambda f)$ la fréquence spatiale de Fourier associée à la position \mathbf{x} dans l'image, l'expression de l'image en fonction de l'amplitude complexe $A_{\lambda,t}(\mathbf{u})$ sur le miroir primaire du télescope s'écrit finalement

$$i_{\lambda,t}(\mathbf{x}) = \frac{1}{(\lambda f)^2} \times |\mathcal{F}[P_{\lambda}(\mathbf{u}) A_{\lambda,t}(\mathbf{u})]|^2. \quad (1.11)$$

1.2 Turbulence Atmosphérique

1.2.1 Description physique de la turbulence

Le phénomène de turbulence correspond aux mouvements aléatoires d'un fluide lorsque celui-ci rencontre des obstacles à une vitesse suffisante. La première description générale et réaliste du mouvement des fluides a été faite par Navier en 1823. Celui-ci a en effet repris le travail d'écriture des équations de Newton pour les particules individuelles d'un fluide, réalisé par Euler, en y ajoutant un terme de friction entre les couches de fluide, proportionnel au coefficient de viscosité ϱ et aux variations v de la vitesse moyenne du fluide. Le travail expérimental de Reynolds a ensuite permis de caractériser la turbulence comme un état succédant à l'état d'écoulement laminaire pour un nombre de Reynold $Re = dv/\varrho$ suffisamment élevé ($\approx 10^3$), où d est le diamètre caractéristique de l'écoulement. Reynolds a également mis en évidence certaines propriétés fondamentales de la turbulence, comme l'augmentation du nombre de petits tourbillons avec le nombre Re .

En 1922 Lewis Fry Richardson formule la première description en cascade radiative des turbulences pleinement développées, c'est à dire des turbulences à grand nombre de Reynolds. Il s'inspire alors en partie d'un poème de Jonathan Swift :

"So, nat'ralists observe, a flea
Hath smaller fleas that on him prey ;
And these have smaller yet to bite 'em,
And so proceed ad infinitum. "

La turbulence développée est de nature fractale : des tourbillons de grande taille transfèrent leur énergie à des tourbillons de plus petite taille et ainsi de suite jusqu'à dissipation de l'énergie par frottement visqueux.

A l'aide d'un argument dimensionnel lié à la conservation de l'énergie et en faisant l'hypothèse que le taux de perte d'énergie est indépendant de la taille l du tourbillon, Kolmogorov a ensuite montré que la variation de la vitesse de l'écoulement turbulent entre deux instants est proportionnelle à la racine cubique de l'intervalle de temps séparant ces deux instants. De façon équivalente, d'après l'ergodicité de la turbulence, la différence de vitesse entre deux points est proportionnelle à la racine cubique de la distance entre les deux points. Cela est décrit par la loi de Kolmogorov [Kolmogorov, 1941] sur la densité spectrale d'énergie turbulente :

$$\Phi_E(\mathbf{u}) \propto \mathbf{u}^{-11/3}, \quad (1.12)$$

qui est valable dans le domaine inertiel $l_0 \ll l \ll L_0$.

l_0 est l'échelle interne de la turbulence. C'est l'échelle à laquelle la viscosité du fluide entre en jeu, ce qui arrive approximativement pour un nombre de Reynolds proche de 1. Pour l'atmosphère terrestre l_0 est de l'ordre du mm.

L_0 est l'échelle externe de la turbulence. Elle correspond à la taille caractéristique du plus gros tourbillon de la couche turbulente. Elle est typiquement de l'ordre de une à quelques dizaines de mètres pour la turbulence atmosphérique [Martin et al., 2000].

En réalité la dimension fractale de dissipation de l'énergie de la turbulence n'est pas tout à fait égale à trois. La statistique dérivée de la loi de Kolmogorov ne rend notamment parfois pas bien compte de l'évolution temporelle de la turbulence.

1.2.2 Description statistique

Les hypothèses et modélisations mathématiques décrivant le phénomène turbulent sont présentées en détail dans Tatarskii [1971]. Je rappelle ici uniquement les hypothèses et les résultats statistiques les plus importants qui sont utilisés dans mon travail, étant entendu que des changements de statistique ne modifient pas les méthodes que j'ai développées.

1.2.2.1 Hypothèses statistiques

La première hypothèse qui est généralement faite est de considérer que les fonctions aléatoires de la turbulence sont localement stationnaires et isotropes dans le domaine inertiel. Ainsi les champs aléatoires ne dépendent localement que de la distance entre deux points du champ. De plus la densité spectrale d'énergie à une dimension est simplement reliée à la densité spectrale d'énergie à trois dimensions par :

$$\Phi_E^{1-D}(f) = -\frac{1}{(2\pi)^3 f} \frac{d\Phi_E^{3-D}(f)}{df}. \quad (1.13)$$

Dans le domaine de Fourier associé on préfère la description statistique par la fonction structure $D_E(\rho) = \langle |E(r + \rho) - E(r)|^2 \rangle$, plutôt que par la fonction de corrélation $B_E(\rho)$ qui est directement la transformée de Fourier de la densité spectrale d'énergie. En effet l'intégration de la densité spectrale d'énergie présente une singularité en 0 dans le cas de loi de Kolmogorov. De plus la fonction structure présente l'avantage d'être également adaptée à la description du cas mi-stationnaire. Avec l'hypothèse isotrope la fonction de structure à une dimension est reliée à la densité spectrale à une dimension par :

$$\begin{aligned} D_E(\rho) &= 2 [B_E(0) - B_E(\rho)] \\ &= 4\pi \int_{-\infty}^{+\infty} [1 - \cos(2\pi f \rho)] \times \Phi_E(f) df. \end{aligned} \quad (1.14)$$

1.2.2.2 Fluctuations spatiales de l'indice de réfraction d'une couche turbulente

Obukhov [1949] et Yaglom [1949] ont montré qu'une quantité physique à l'intérieur d'un écoulement turbulent, suit également la loi de Kolmogorov si elle n'affecte

pas l'écoulement turbulent (quantité passive), et si elle est indépendante des réactions chimiques qui ont lieu dans l'écoulement (quantité conservatrice). C'est notamment le cas pour l'air atmosphérique des paramètres macroscopiques de température, de pression et de concentration en vapeur d'eau. De même les fluctuations de l'indice de réfraction $n(\mathbf{r}, h_i, t)$ dans une couche turbulente à l'altitude h_i dépendent de ces derniers paramètres macroscopiques, et l'on montre [Roddier, 1981] que leur densité spectrale d'énergie à trois dimensions suit la loi de Kolmogorov :

$$\Phi_n(f) = \frac{C_N^2}{103.182} \times f^{-11/3}, \quad (1.15)$$

où C_N^2 caractérise la force des fluctuations d'indice de la couche turbulente et est appelée constante de la fonction de structure d'indice :

$$D_n(\rho) = C_N^2 \times \rho^{2/3}. \quad (1.16)$$

Ces expressions ne sont valables que dans le domaine inertiel. En approximant la décroissance rapide de la turbulence pour $\rho < l_0$ et en modifiant la densité spectrale de Kolmogorov pour la rendre intégrable à l'origine, Tatarskii a proposé le spectre de Von Karman de la turbulence :

$$\Phi_n(f) = \frac{C_N^2}{103.182} \times \frac{(f^2 + f_m^2)^{-11/6}}{\exp[(f/f_M)^2]}, \quad (1.17)$$

où $f_M = 1/l_0$ et $f_m = 1/L_0$. Le défaut du spectre de Von Karman est de ne pas bien prendre en compte les échelles $f < f_m$. D'autres expressions ont été proposées pour mieux décrire les fluctuations de l'indice de réfraction. Elles sont comparées dans [Voitsekhovich, 1995]. Je me limite pour la suite de mon travail à des spectres et covariances dérivés de la loi de Kolmogorov.

1.2.2.3 Fluctuations spatiales de la phase sur la pupille

On considère en général pour l'observation astronomique, que la turbulence atmosphérique n'entraîne pas d'absorption ni de diffusion hétérogène de la lumière, ce qui peut toutefois être remis en cause en fonction du site d'observation, du diamètre du télescope et également des conditions particulières d'observation. L'atmosphère est alors modélisée comme une stratification de couches turbulentes d'épaisseurs négligeables.

La diffraction de Fresnel par une mince couche turbulente à l'altitude h_i entraîne en principe une répartition hétérogène de l'amplitude dans le plan du miroir primaire au sol. Ces fluctuations se manifestent dans le temps par de la scintillation (clignotement) dans les images. Si l'on observe ce qui se passe sur la pupille du télescope on observe alors des ombres volantes, comme on peut également en observer de jour avec la réflexion de la lumière du soleil au fond d'une piscine. L'importance des fluctuations d'amplitude à cause de la diffraction de Fresnel croît avec l'altitude h séparant

la pupille et la couche turbulente et également avec le C_N^2 . Les fluctuations deviennent significatives lorsque la taille de la première zone de Fresnel est de l'ordre de la taille de cohérence de l'atmosphère, c'est à dire lorsque $h \approx r_0^2/\lambda$, où r_0 est le diamètre caractéristique de cohérence de la turbulence par rapport aux fluctuations de la phase, et est également appelé le diamètre de Fried. C'est le cas par exemple pour des couches turbulentes situées au-delà de $h=45$ km, avec un $r_0=15$ cm à $\lambda=500$ nm.

Les études de profil de C_N^2 avec l'altitude ont montré que les couches turbulentes sont généralement regroupées à de plus basses altitudes [Tokovinin et al., 2002]. On fait donc souvent l'hypothèse que l'amplitude de l'onde EM au sol ne dépend pas de la position \mathbf{r} sur la pupille (hypothèse de champ proche), pour les télescopes situés sur les tout meilleurs sites d'observation. On considère alors que la traversée de l'atmosphère entraîne seulement un retard de phase. Celui-ci est hétérogène sur la pupille et s'obtient par la somme des retards induits par les différentes couches traversées :

$$\varphi_\lambda(\mathbf{r}) = \frac{2\pi}{\lambda} \int_{h_0}^{h_{\max}} n_{\lambda,t}(\mathbf{c}) d\mathbf{c}. \quad (1.18)$$

En combinant l'équation de phase (1.18) avec l'équation de la fonction de structure de l'indice de réfraction (1.16), Tatarskii [1961] obtient la fonction de structure de la phase :

$$D_\varphi(\rho) = 2.91 \left(\frac{2\pi}{\lambda} \right)^2 \cos(\gamma)^{-1} \rho^{5/3} \int_{h_0}^{h_{\max}} C_N^2(h) dh, \quad (1.19)$$

où γ désigne l'angle entre la visée et le zénith. On remarque que cette expression ne prend pas en compte la chromaticité du chemin suivi par la lumière.

1.2.2.4 Fluctuations temporelles de la phase sur la pupille

Les caractéristiques temporelles de la turbulence sont reliées aux caractéristiques spatiales par l'hypothèse de Taylor de la turbulence gelée [Taylor, 1938]. Cette hypothèse consiste à décrire l'évolution temporelle d'une couche turbulente à l'altitude h , uniquement par une translation à la vitesse $\mathbf{v}(h)$. La variation temporelle du retard de phase à l'altitude h est ainsi reliée à sa variation spatiale par :

$$\varphi_h(\mathbf{r}, t + \tau) = \varphi_h(\mathbf{r} - \tau \mathbf{v}(h), t). \quad (1.20)$$

On en déduit l'expression de la fonction de structure temporelle de la phase :

$$D_\varphi(\tau) = 2.91 \left(\frac{2\pi}{\lambda} \right)^2 \cos(\gamma)^{-1} (\tau \bar{v})^{5/3} \int C_N^2(h) dh, \quad (1.21)$$

où $\bar{v} = \left(\frac{\int v(h)^{5/3} C_N^2(h) dh}{\int C_N^2(h) dh} \right)^{3/5}$ est la vitesse caractéristique de la turbulence.

1.3 Effets de la turbulence dans les images

1.3.1 Effets dans les images longues poses

1.3.1.1 Fonction de transfert longue pose

Les fluctuations spatio-temporelles de la phase sur la pupille du télescope entraînent une perte de détail dans les images acquises en continu au foyer du télescope (images *longues poses*), par rapport à la résolution angulaire théorique du télescope $\alpha_c = \lambda/D$. Cette perte peut-être quantifiée dans l'espace de Fourier associé aux images, comme on l'a vu à la section 1.1.1, à l'aide de la fonction de transfert longue pose du système d'imagerie :

$$\begin{aligned} \langle H(\mathbf{u}) \rangle &= \frac{\left\langle \iint P(\mathbf{u}' + \mathbf{u}) e^{j\varphi(\mathbf{u}' + \mathbf{u})} P(\mathbf{u}') e^{-j\varphi(\mathbf{u}')} d\mathbf{u}' \right\rangle}{\iint |P(\mathbf{u}')|^2 d\mathbf{u}'} \\ &= T(\mathbf{u}) \times B(\mathbf{u}), \end{aligned} \quad (1.22)$$

où $B(\mathbf{u}) = \langle e^{j[\varphi(\mathbf{u}' + \mathbf{u}) - \varphi(\mathbf{u}')] } \rangle$ est la fonction de transfert longue pose de l'atmosphère et $\langle . \rangle$ représente un moyennage sur t . En supposant que l'amplitude de l'onde ne fluctue pas spatialement sur la pupille, et comme $\varphi(\mathbf{u}' + \mathbf{u}) - \varphi(\mathbf{u}')$ est une variable gaussienne on montre [Fried, 1966] que $B(\mathbf{u})$ s'écrit :

$$B(\mathbf{u}) = \exp \left[-\frac{1}{2} \langle (\varphi(\mathbf{u}' + \mathbf{u}) - \varphi(\mathbf{u}'))^2 \rangle \right]. \quad (1.23)$$

L'ergodicité de la phase qui découle de l'hypothèse de Taylor permet ensuite de remplacer la moyenne temporelle de l'équation (1.23) par une moyenne spatiale et de réexprimer la fonction de transfert longue pose en fonction de la fonction de structure des fluctuations spatiales de la phase :

$$B(\mathbf{u}) = \exp \left[-\frac{1}{2} D_\varphi(\mathbf{u}) \right]. \quad (1.24)$$

1.3.1.2 Paramètre de Fried

Fried [1966] a introduit le paramètre r_0 pour exprimer les dégradations optiques dues à la turbulence dans la fonction de transfert $\langle B(u) \rangle$. Comme on l'a vu à la section 1.1.1 la résolution des images formées par un télescope augmente avec le diamètre physique D du télescope. Le paramètre de Fried r_0 est donc défini comme le diamètre équivalent en terme de résolution auquel la turbulence limite le télescope dans les images longues poses, si l'on ne corrige pas les effets délétères de la turbulence :

$$\int_{u=0}^{u=r_0} T(u) du = \int_{u=0}^{u=+\infty} \langle B(u) \rangle du \quad (1.25)$$

Dans le cas d'une pupille circulaire, Fried obtient :

$$r_0(\gamma, \lambda) = \left[0.423 \times \left(\frac{2\pi}{\lambda} \right)^2 \cos(\gamma)^{-1} \int C_N^2(h) dh \right]^{-3/5}. \quad (1.26)$$

Fried montre également dans cet article que r_0 caractérise les fluctuations spatiales de la phase de l'onde EM sur la pupille : r_0 est la taille moyenne pour laquelle la phase fluctue en variance de 1 radian². r_0 vaut typiquement une quinzaine de cm à 500 nm, au zénith, pour les meilleurs sites d'observation.

1.3.1.3 Résolution angulaire longue pose

En combinant les équations (1.26) et (1.19), la fonction structure des fluctuations spatiales de la phase s'écrit simplement en fonction de r_0 dans le modèle de Kolmogorov :

$$D_\varphi(\rho) = 6.88 \left(\frac{\rho}{r_0} \right)^{5/3}. \quad (1.27)$$

De même par l'équation (1.24) on obtient l'expression de la fonction de transfert longue pose de l'atmosphère en fonction du r_0 :

$$B(u) = \exp \left[-3.44 \times \left(\frac{u}{r_0/\lambda} \right)^{5/3} \right]. \quad (1.28)$$

La transmission atmosphérique des fréquences spatiales exprimée dans (1.28) est quasiment nulle au-delà de la fréquence de coupure $u_{\text{long}} = r_0/\lambda$. On définit donc généralement la résolution angulaire des images longues poses limitées par la turbulence des grands télescopes ($D \gg r_0$) par :

$$\alpha_{\text{long}} = \frac{\lambda}{r_0}, \quad (1.29)$$

lorsqu'on ne corrige pas les effets de la turbulence. α_{long} est aussi communément appelé le *seeing* de la turbulence. Il vaut 0.69 secondes d'arc pour un r_0 typique de 15 cm à 500nm, et 4.74 secondes d'arc pour le r_0 correspondant à 2.5 μm dans l'infrarouge.

1.3.2 Effets dans les images courtes poses

1.3.2.1 Images tavelées

L'interférence des différentes zones de cohérence de la phase de l'onde EM — de diamètre r_0 chacune sur la pupille —, crée des noeuds d'interférence ou *tavelures* dans les images courtes poses au foyer du télescope. Ces tavelures sont en nombre de l'ordre de $(D/r_0)^2$ et ont chacune à peu près la taille de la tache d'Airy $\simeq \lambda/D$. On observe ces

tavelures pour des télescopes de diamètre $D \gg r_0$ et pour des temps de pose maximum de l'ordre de quelques ms voir quelques dizaines de ms dans le visible.

La dégradation de la qualité optique des images courtes poses se traduit par une atténuation des hautes fréquences spatiales dans la fonction de transfert courte pose $\langle |H_c(u)| \rangle$. Toutefois les hautes fréquences spatiales restent présentes dans les images et peuvent servir pour obtenir des images longues poses à haute résolution angulaire, ce qui est discuté plus en détail à la section 1.4.

1.3.2.2 Cohérence temporelle

Le temps de cohérence τ_0 caractérise les corrélations temporelles du front d'onde et la vitesse d'évolution des dégradations optiques induites par la turbulence dans les images. Il peut être défini à l'analogie du r_0 comme le temps pour lequel la fluctuation temporelle du front d'onde en variance vaut 1 radian². En combinant la définition du paramètre de Fried de l'équation (1.26) et l'équation de la fonction structure des fluctuations temporelles de la phase (1.21), on obtient :

$$D_\varphi(\tau) = 6.88 \left(\frac{\bar{v}\tau}{r_0} \right)^{5/3}, \quad (1.30)$$

qui donne pour 1 radian² l'expression de τ_0 en fonction de r_0 et de la vitesse caractéristique de la turbulence \bar{v} :

$$\tau_0 = 0.314 \times \frac{r_0}{\bar{v}}. \quad (1.31)$$

L'ajustement expérimental de τ_0 [Karo and Schneiderman, 1978; Ebersberger and Weigelt, 1985] montre que τ_0 vaut typiquement de l'ordre de 5 ms à 500 nm dans le visible.

Divers temps caractéristiques sont exprimés en fonction de τ_0 pour les techniques visant à obtenir de hautes résolutions angulaires. Walker montre par exemple [Walker, 1978], que le temps de pose $T_0 = 2 \times \tau_0$ maximise le rapport signal pour l'interférométrie des tavelures.

1.3.2.3 Fonction de transfert courte pose

Comme la fonction de transfert courte pose $H_c(u)$ est une fonction aléatoire, on a recours à la transmission moyenne de puissance $\langle |H_c(u)|^2 \rangle$ pour décrire le transfert des fréquences spatiales de l'objet dans l'image :

$$\langle |I(u)|^2 \rangle = \langle |H_c(u)|^2 \rangle \times |O(u)|^2. \quad (1.32)$$

Le modèle de $\langle |H_c(u)|^2 \rangle$ le plus généralement utilisé pour les télescopes de diamètre D grand devant r_0 , est le modèle log-normal de Korff [1973]. Il s'écrit comme la somme de deux termes asymptotiques :

$$\langle |H_c(u)|^2 \rangle = \langle |H_c(u)|^2 \rangle_{\text{BF}} + \langle |H_c(u)|^2 \rangle_{\text{HF}}, \quad (1.33)$$

qui valent respectivement :

$$\langle |H_c(u)|^2 \rangle_{\text{BF}} = \exp \left[-6.88 \times \left(\frac{u}{r_0/\lambda} \right)^{5/3} \times \left(1 - \frac{u}{D/\lambda} \right)^{1/3} \right] \times |T(u)|^2 \quad \text{pour } u \ll \frac{r_0}{\lambda} \quad (1.34)$$

$$\langle |H_c(u)|^2 \rangle_{\text{HF}} = 0.435 \times \left(\frac{r_0}{D} \right)^2 \times T'^2(u) \quad \text{pour } u \gg \frac{r_0}{\lambda} \quad (1.35)$$

où $T'^2(u) = \int |P(u') P(u' + u)|^2 du' / \int |P(u')|^4 du'$.

Le terme basses fréquences correspond à peu près à la fonction de transfert longue pose lorsque les images longues poses sont la somme des images courtes poses recen-
trées sur leur barycentre.

Dans le terme hautes fréquences, les fréquences spatiales des objets observés sont uniformément atténuées par rapport à la transmission $T'^2(u)$.

1.3.3 Illustration

Pour illustrer l'effet de la turbulence en imagerie astronomique, je montre des images de sources ponctuelles et leurs fonctions de transfert, que j'ai acquises au début de ma thèse à l'Observatoire de Haute-Provence, grâce au montage ATTILA (cf. thèse de Julien Girard [2005]).

À gauche sur la figure 1.2, on voit une image courte pose typique avec sa structure tavelée (les speckles). La longueur d'onde de l'acquisition est de 700 nm et le temps de pose est de 40 ms, ce qui est légèrement supérieur au temps de cohérence de la turbulence sur le site de l'OHP, à cette longueur d'onde. Cela se traduit par une perte de contraste des tavelures. Les acquisitions en courtes poses transmettent néanmoins des hautes fréquences spatiales par rapport aux acquisitions en longues poses, comme on le voit sur les fonctions de transfert en puissance ($\langle |H_c(\mathbf{u})|^2 \rangle$ et $H(\mathbf{u})^2$) représentées sur la figure 1.3, pour deux longueurs d'onde différentes, 700 et 570 nm). On estime à gauche sur la figure 1.4, la perte de résolution angulaire causée par la turbulence pour l'observation d'objets résolus en longues poses. En effet, cette dégradation correspond au rapport des largeurs à mi-hauteur entre les taches images de sources non résolues, avec et sans turbulence. Ce rapport correspond également au D/r_0 et il est de l'ordre de 20 dans le cas présent, à 700 nm. La largeur de la réponse impulsionnelle instrumentale (télescope seul) sur le CCD a été calibrée grâce au recoupement de la séparation angulaire du système d'étoile double Castor, calculée ce jour-là, et de la distance entre les deux étoiles mesurée dans l'autocorrélation moyenne des images courtes poses (à droite sur la figure 1.4). Je trouve en effet d'après ce recoupement, que 1 pixel CCD correspond à 47 mas, soit environ la moitié de la largeur à mi-hauteur théorique du 1.52 m de l'OHP à 700 nm ($0.98\lambda/D \simeq 93$ mas). La valeur angulaire d'un pixel CCD

que j'ai calculée sur Castor, correspond à la valeur prévue d'après les spécifications du montage optique ATTILA (48 mas donné par Julien Girard dans sa thèse).

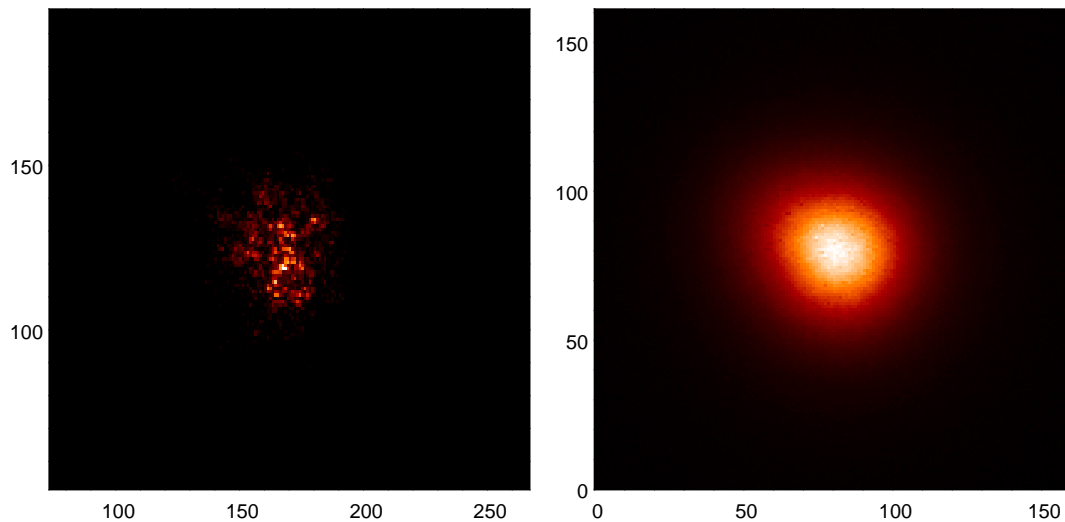


FIG. 1.2 – Image courte pose (à gauche) et image longue pose (à droite) acquises à l'observatoire de Haute-Provence. On observe l'élargissement de la réponse impulsionnelle longue pose causé par la turbulence, et la structure tavelée de l'image courte pose qui permet la transmission de hautes fréquences spatiales.

1.4 Techniques pour obtenir les Hautes Résolutions Angulaires dans les images longues poses

Dans cette section je ne présente que les principales techniques de HRA avec un seul télescope. En réalité ces techniques comprennent aussi l'interférométrie à plusieurs télescopes, qui est en plein essor, et pour laquelle on peut par exemple se reporter à [Lagrange et al., 1997; Saha, 2002].

1.4.1 Techniques de HRA a posteriori

On parle de techniques a posteriori ou d'imagerie des tavelures (speckles), pour désigner les techniques où la réponse impulsionnelle (psf) de l'atmosphère est déconvoluée des images courtes poses, afin d'obtenir les hautes résolutions angulaires des télescopes de diamètre $D \gg r_0$. Les débuts des techniques a posteriori remontent aux travaux de Labeyrie en interférométrie des tavelures [Labeyrie, 1970]. En utilisant la relation de filtrage de l'équation (1.3) et en s'appuyant sur la présence des hautes fréquences spatiales dans les images courtes poses, Labeyrie a en effet montré que le

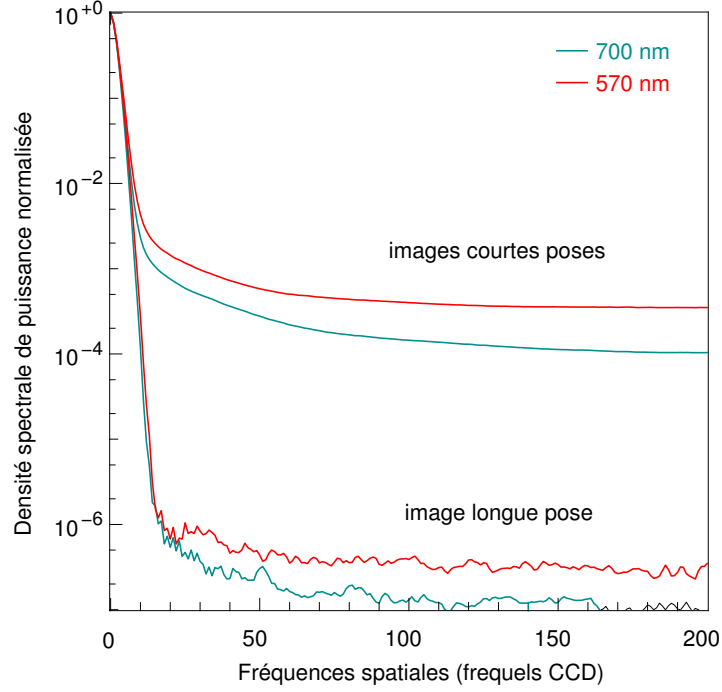


FIG. 1.3 – Fonctions de transfert (en puissance) des images courtes poses et longue pose, mesurées simultanément pour deux longueurs d’onde du visible avec le montage ATTILA, à partir d’une étoile non résolue.

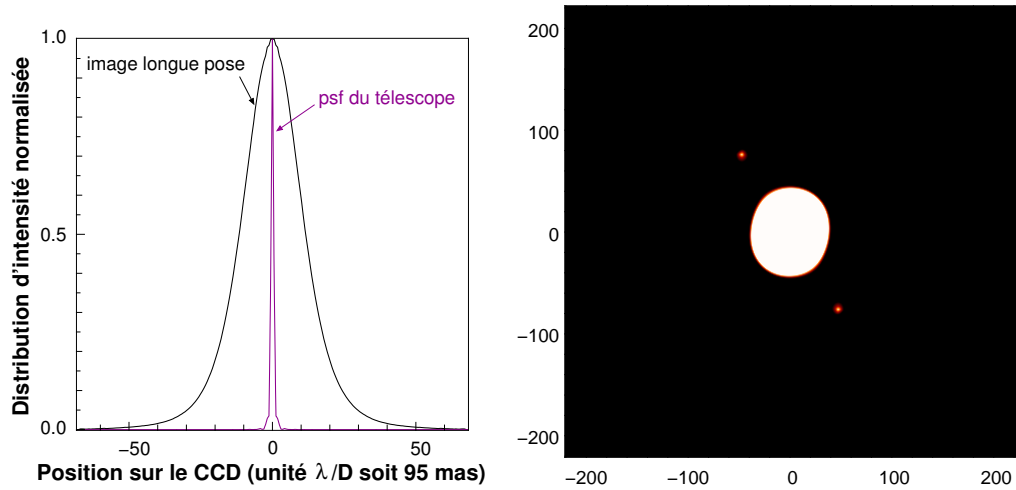


FIG. 1.4 – A gauche, comparaison entre le profil radial de la psf longue pose avec turbulence, et celui de la psf instrumentale sans turbulence. Le rapport des largeurs à mi-hauteur vaut $D/r_0 \approx 20$ et correspond à la perte de résolution angulaire causé par la turbulence. A droite, autocorrélation moyenne des images courtes poses d’un système d’étoile double (Castor).

spectre de puissance de l'objet $|O(\mathbf{u})|^2$ peut-être obtenu par intégration des spectres de puissance des images courtes poses :

$$|\widetilde{O}(\mathbf{u})|^2 = \frac{\sum_{i=1}^{i=n} |I_i(\mathbf{u})|^2 / n}{\langle |H_c(\mathbf{u})|^2 \rangle}, \quad (1.36)$$

où $|I_i(\mathbf{u})|^2$ désigne le spectre de puissance de la i -ème image courte pose, et $\langle |H_c(\mathbf{u})|^2 \rangle$ est le spectre de puissance courte pose du système d'imagerie atmosphère+télescope, considéré aux fréquences spatiales \mathbf{u} où la transmission n'est pas nulle.

$\langle |H_c(\mathbf{u})|^2 \rangle$ est calibré en moyennant les spectres de puissance des images courtes poses d'une source ponctuelle. Cette source de référence doit être suffisamment proche de l'objet observé, et la calibration doit être effectuée dans un temps suffisamment proche des acquisitions de l'objet, pour que le r_0 n'ait pas varié. Cela est discuté plus en détail dans Perrier [1989]. Puisque la phase de $O(\mathbf{u})$ n'est pas déterminée, l'interférométrie des tavelures de Labeyrie ne permet de retrouver que l'autocorrélation de l'objet. Elle ne sert donc que pour l'étude d'objets astronomiques simples comme des étoiles doubles et des étoiles triples, ou des disques circulaires ou aplatis sans structure fine, ou des combinaisons simples de composantes élémentaires de ce type.

Les méthodes de Knox-Thompson [Knox and Thompson, 1974] et du bispectre [Weigelt, 1977] ont ensuite été proposées pour retrouver la phase de $O(\mathbf{u})$, et ainsi obtenir l'objet avec le maximum de détails. Ces méthodes s'appuient sur la présence de l'information de phase à haute fréquence dans les expressions respectives :

$$I_i^{KT}(\mathbf{u}, \mathbf{v}) = \langle I_i(\mathbf{u}) \times I_i^*(\mathbf{u} + \mathbf{v}) \rangle \quad (1.37)$$

$$I_i^{BS}(\mathbf{u}, \mathbf{v}) = \langle I_i(\mathbf{u}) \times I_i(\mathbf{v}) \times I_i^*(\mathbf{u} + \mathbf{v}) \rangle. \quad (1.38)$$

Ces méthodes nécessitent cependant de stocker un volume important de données. Elles sont décrites plus en détail dans [Thiébaud, 1994; Ayers et al., 1988].

Les méthodes itératives de reconstruction de phase de Fourier [Fienup and Feldkamp, 1980] puis de déconvolution aveugle par maximum de vraisemblance [Lane, 1992] ont ensuite été proposées pour l'imagerie astronomique instantanée. Toutefois elles requièrent souvent une bonne connaissance de départ de la phase et de l'objet. Les méthodes d'optimisation globale que j'ai développées en reconstruction de phase et que je présente au chapitre 3 permettent dans ce sens de traiter des conditions de turbulence relatives (D/r_0) et de rapport signal à bruit plus difficiles.

1.4.2 Optique Adaptative

Les techniques de HRA temps-réel avec une Optique Adaptative (OA) permettent d'intégrer au foyer d'un télescope des images courtes poses, dont les dégradations optiques dues à l'atmosphère terrestre ont été partiellement corrigées (fig. 1.5). Ces techniques peuvent être combinées à des corrections a posteriori des images longues poses

obtenues, lorsqu'on souhaite obtenir une résolution angulaire encore plus proche de la résolution limite du télescope. La technique d'Optique Adaptative a été proposée par Babcock [1953] et a tout d'abord été développée pour des applications militaires. Elle est ensuite devenue la solution privilégiée de la HRA pour l'imagerie astronomique longue pose basée sur la Terre. L'OA a fait l'objet de développements considérables à la fin du XX^{ième} siècle, en dépit de contraintes technologiques et physiques lourdes. Grâce à cet essor, l'OA est aujourd'hui utilisée pour tout type d'application physique où il faut estimer des aberrations optiques. Elle est devenue un véritable *business* bien que des techniques a posteriori puissent dans certains cas la remplacer à moindre frais.

1.4.2.1 Corrections et estimations du front d'onde en temps réel

La compensation des aberrations de phase (front d'onde) est effectuée dans des plans conjugués au miroir primaire du télescope. Un grand miroir basculable sert tout d'abord à corriger les pentes globales de la phase (modes de tip-tilt), qui dominent les fluctuations spatio-temporelles de la phase. Des petits miroirs piezo-électriques déformables sont utilisés dans un second plan pour corriger les aberrations du front d'onde d'ordres plus élevés que la pente globale. Ces petits miroirs sont en nombre de l'ordre de $(D/r_0(\lambda))^2$ (pour l'instant les OA sur les grands télescopes astronomiques ont un nombre d'actuateurs correspondant à un échantillonnage adéquat pour une correction dans le proche infra-rouge principalement), et doivent effectuer des pas de déformation avec une précision bien meilleure que le micromètre et à une cadence supérieure au kHz. La séparation des modes de tip-tilt permet ainsi de considérablement simplifier la dynamique des petits miroirs.

La compensation appliquée par les miroirs correcteurs est calculée à l'aide des mesures d'un senseur de front d'onde. Dans les systèmes actuels d'OA le senseur est placé après les miroirs correcteurs. Il mesure ainsi les aberrations résiduelles du front d'onde et asservit en boucle fermée les miroirs, au fur et à mesure que le front d'onde évolue. Une revue approfondie des divers types de senseurs de front d'onde et de leurs performances est accomplie dans Rousset [1999].

Le senseur de front d'onde le plus répandu est le senseur linéaire de Shack - Hartmann [Platt and Shack, 1971]. Il divise spatialement le front d'onde en sous-faisceaux, qui sont focalisés par un réseau de micro-lentilles. Le centroïde de chaque sous-faisceau focalisé est un bon estimateur de la pente locale du front d'onde, lorsque la référence lumineuse est une source non résolue pour une sous pupille (sinon il faut relier ce centroïde au gradient moyen et non pas à la pente locale de la surface d'onde), et pourvu que les aberrations locales d'ordres plus élevés ne soient pas trop importantes. Le front d'onde peut alors être reconstruit en utilisant les estimations de pentes locales et de bons a priori. L'avantage de ce senseur géométrique est que les centroides ne dépendent pas de la longueur d'onde. Il peut donc fonctionner en mode achromatique pour intégrer plus de lumière. De plus il peut également fonctionner avec

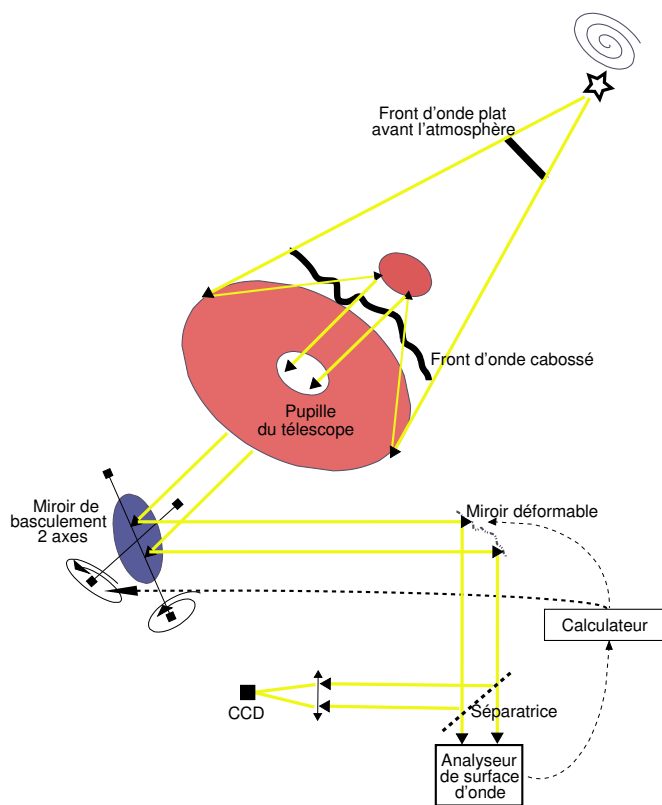


FIG. 1.5 – Principe de fonctionnement d'une optique adaptative.

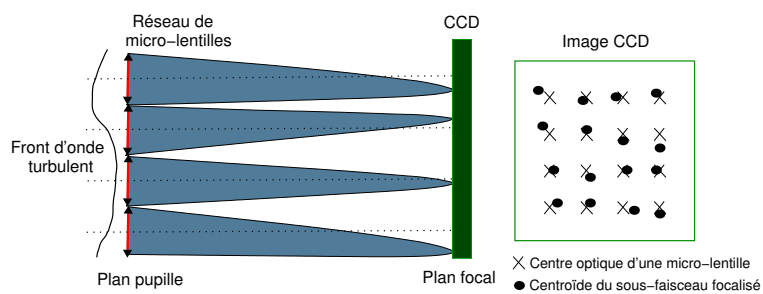


FIG. 1.6 – Principe de fonctionnement de l'analyseur de surface d'onde de Shack-Hartmann.

des sources étendues, en utilisant des corrélations entre les différentes imagerie de la source pour estimer les pentes locales [Michau et al., 1992].

Un senseur mesurant la courbure de la phase, d'après les sur-intensités et sous-intensités qui sont associées à la courbure dans des plans situés de part et d'autre du plan focal, a été proposé par Roddier [Roddier, 1988]. Son utilisation se limite toutefois à l'estimation d'un nombre moins important de modes que le Shack-Hartmann, c'est à dire pour des faibles D/r_0 . Il sert notamment à mesurer les aberrations statiques des télescopes. D'autres senseurs prometteurs, comme le senseur pyramidal [Ragazzoni, 1996] ont également été proposés pour l'Optique Adaptative. Toutefois le Shack-Hartmann bénéficie d'une plus grande simplicité et d'une plus grande maîtrise.

Les méthodes de reconstruction de phase à partir d'images tavelées focalisées peuvent également servir à l'analyse de surface d'onde pour l'Optique Adaptative. Elles peuvent être envisagées à l'analogie d'un Shack-Hartmann pour estimer les aberrations d'ordres plus élevés que les pentes des sous-pupilles [Cannon, 1995; van Dam and Lane, 2002a]. Dans ce sens la méthode de reconstruction de phase à partir d'une image tavelée que j'ai développée, permet d'envisager un système d'optique adaptative en boucle ouverte avec des rapports d/r_0 élevés, où d est le diamètre d'une sous-pupille. Ces méthodes s'appuient cependant sur des images monochromatiques et nécessitent donc un flux de référence monochromatique élevé. La diversité polychromatique de phase peut permettre d'améliorer le flux en utilisant des images à plusieurs longueurs d'onde pour estimer le même chemin optique. La diversité polychromatique permet également, comme je le montre au chapitre 5, de reconstruire des phases pour des D/r_0 plus élevés par rapport au cas monochromatique. Les méthodes de reconstruction de phase ne sont à l'heure actuelle pas encore utilisées pour les corrections en temps réel de l'optique adaptative, car elles requièrent une puissance de calcul trop élevée. Elles sont toutefois très attrayantes par la simplicité de l'instrumentation mise en jeu et puisqu'elles permettent d'estimer toutes les aberrations jusqu'au plan focal des images.

1.4.3 Source de référence pour les techniques HRA

1.4.3.1 Angle d'isoplanétisme

Pour estimer les aberrations optiques dues à la turbulence avec suffisamment de précision, il faut disposer d'un flux de référence relativement élevé, et ce quelle que soit la technique a posteriori ou le senseur de front d'onde utilisé. Dans le cas de l'observation astronomique la lumière est particulièrement raréfiée et la turbulence évolue rapidement. La plupart du temps il n'est donc pas envisageable d'utiliser la lumière des objets astronomiques observés pour les corrections. Il faut alors trouver une source de référence brillante, de préférence ponctuelle, située dans un domaine angulaire suffisamment proche de l'objet observé (domaine d'isoplanétisme).

Comme on vient de le voir ce sont principalement les aberrations de phase causées par l'atmosphère qui sont responsables des dégradations des images astronomiques pour les télescopes situés sur les tout meilleurs sites terrestres d'observation. Le domaine d'isoplanétisme est donc en pratique fixé par la turbulence. L'angle d'isoplanétisme est défini par Fried à l'analogie des paramètres r_0 et τ_0 comme l'angle d'observation de deux sources ponctuelles, pour lequel les aberrations de phase induites par l'atmosphère diffèrent en variance de 1 rad². En utilisant l'expression (1.26) du r_0 , il vient [Fried, 1982] :

$$\theta_0 = 0.314 \times \frac{r_0}{\bar{h}}, \quad (1.39)$$

où \bar{h} est l'altitude caractéristique de la turbulence :

$$\bar{h} = \left(\frac{\int h^{5/3} C_n^2(h) dh}{\int C_n^2(h) dh} \right)^{3/5}. \quad (1.40)$$

Comme on le voit dans la définition de \bar{h} et sur la figure 1.7, ce sont les couches turbulentes situées aux altitudes les plus élevées, pour des forces de fluctuation d'indice $C_n^2(h)$ comparables, qui déterminent l'angle d'isoplanétisme.

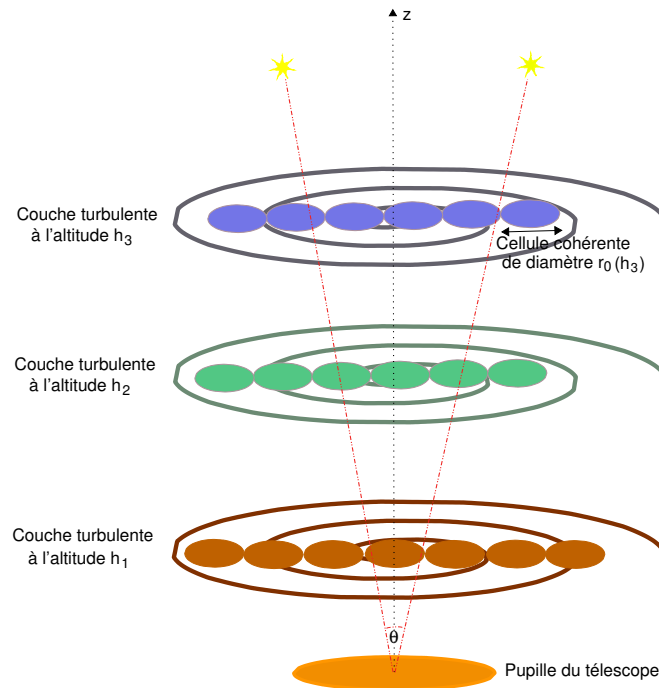


FIG. 1.7 – Anisoplanétisme entre deux objets observés.

Tout comme les autres paramètres atmosphériques de Fried, l'angle d'isoplanétisme est chromatique et varie comme $\lambda^{1.2}$. Par exemple pour une altitude caractéristique de $\bar{h} = 5$ km, l'angle d'isoplanétisme vaut ~ 2 secondes d'arc avec un r_0 typique de 15 cm à 500 nm dans le visible, tandis qu'il vaut plus de 13 secondes d'arc pour le r_0 équivalent à $2.5 \mu\text{m}$ dans l'infrarouge.

Ces valeurs sont plus petites que le champ d'observation des télescopes de grand diamètre. De plus la possibilité de trouver une source de calibration dans la partie spécifique du champ pour laquelle on souhaite obtenir les hautes résolutions angulaires, décroît rapidement pour les courtes longueurs d'onde. En pratique on montre que la proportion de couverture du ciel pour effectuer des corrections à l'aide de l'optique adaptative est inférieure à 1% dans le visible [Foy, 2005].

1.4.3.2 Étoile laser artificielle

Le concept d'étoile artificielle laser a été proposé par Foy and Labeyrie [1985] pour remédier au problème d'anisoplanétisme en astronomie et à l'insuffisante couverture du ciel en étoiles de référence suffisamment brillantes. Une source lumineuse de référence est créée dans l'atmosphère terrestre par stimulation laser, à une altitude suffisante par rapport aux couches turbulentes les plus élevées ($h > 30$ km). Le procédé le plus utilisé est celui de l'excitation résonnante des atomes de sodium. Ces atomes sont situés dans une couche de la mésosphère de l'ordre de 10 km autour de $h \approx 95$ km d'altitude. La transition énergétique des électrons du sodium qui est utilisée correspond à une longueur d'onde d'absorption et de réémission spontanée de 589.2 nm.

A cause du rendement modeste de l'ordre de 2% de l'absorption pour cette transition du sodium, du nombre limité d'atomes de sodium dans la section utile de la mésosphère et de leur temps de désexcitation significatif, la première limitation de l'étoile laser est que le flux retourné est parfois insuffisant. Des procédés d'excitation plus efficaces, en utilisant des états atomiques non quantifiés font aujourd'hui l'objet de recherches prometteuses, mais ces procédés ne sont pas encore assez mûrs pour être utilisés pour l'étoile laser.

Le deuxième problème majeur de l'étoile laser est l'effet de cône [Foy and Labeyrie, 1985; Tallon and Foy, 1990]. A cause de la distance finie de l'étoile laser au télescope, les portions de couches turbulentes traversées par la lumière de l'étoile artificielle ne sont pas les mêmes que celles traversées par la lumière provenant de l'objet astronomique. Les systèmes d'optique adaptative multi-conjugués ont été proposés par Foy et Tallon pour résoudre ce problème pour les télescopes de grand diamètre à l'étude, car l'effet de cône augmente avec le diamètre D . Le principe de ces systèmes est de disposer les étoiles lasers de telle sorte, que chaque portion significative de couche turbulente traversée par la lumière d'observation soit caractérisée. Cette méthode est donc une forme de cartographie en 3 dimensions de la turbulence (tomographie). Elle permet également si on applique simultanément des corrections séparées à divers parties

du champ du télescope, d'augmenter la portion de champ correctement corrigée.

Le dernier problème de l'étoile laser artificielle est l'indétermination de la pente de la surface d'onde (tip-tilt) ou de façon équivalente l'indétermination de l'angle d'arrivée de la surface d'onde. Ces deux grandeurs sont différentes, ce qui est discuté plus en détail au chapitre 5, mais elles représentent toutes les deux un déplacement exclusivement, et le déplacement ne peut pas être estimé à partir de l'étoile laser artificielle monochromatique. En effet le décalage global de la lumière est subit également à l'aller du laser et au retour de la lumière réémise à la même longueur d'onde par le sodium, en vertu du principe de retour inverse de la lumière, et puisque le temps de propagation dans l'atmosphère est court devant l'évolution de la turbulence. De façon alternative le faisceau laser peut être émis par un télescope auxiliaire pour que les deux composantes ne s'annulent pas complètement, mais le problème d'indétermination reste intact car on ne sait pas séparer les composantes de déplacement liées à l'aller et au retour. Comme on ne peut pas déterminer dans quelle région du ciel est située l'étoile laser, on ne peut pas estimer quel est le déplacement induit par la turbulence sur l'objet qu'on veut observer. Cette indétermination est critique car 85% de la variance de la surface d'onde atmosphérique, qui doit être estimée pour obtenir les hautes résolutions angulaires, est due à ce déplacement. On peut utiliser des étoiles naturelles situées dans le champ de l'observation pour estimer le déplacement de l'objet. Mais la possibilité de trouver des sources suffisamment brillantes et donc l'efficacité de cette méthode dépend beaucoup de la portion du ciel qui est observée. Le concept d'Etoile Laser Polychromatique a été proposé par Foy [1995] pour obtenir une estimation de la pente de la surface d'onde atmosphérique dans n'importe quelle direction d'observation. Il est actuellement en phase de démonstration pour l'Optique Adaptative dans le visible.

1.4.4 Etoile laser polychromatique

1.4.4.1 Dépendance chromatique du chemin optique

Le principe de mesure de la pente de la surface d'onde avec étoile laser polychromatique repose sur la dépendance chromatique du chemin optique emprunté par la lumière. En effet le temps d'interaction de la lumière avec la matière et donc également la différence $\Delta v = v - v_0$ entre la vitesse v de propagation dans le milieu matériel et la vitesse v_0 de propagation dans le vide sont chromatiques. Il en est donc également de même pour l'indice de réfraction relatif du milieu $n - 1 = \Delta v / v_0$. Dans le cas des gaz parfaits, qui approxime bien l'air atmosphérique, on montre que l'indice de réfraction relatif ne dépend que de la pression P , de la température T et de la longueur d'onde λ , et qu'il s'écrit sous une forme séparable [Edlén, 1966] :

$$n(\lambda, P, T) - 1 = f(\lambda) \times g(P, T) \quad (1.41)$$

Le chemin optique au sol s'écrit donc :

$$\begin{aligned}\delta(\lambda, \mathbf{r}) &= \int_{C(\mathbf{r})} n(\lambda, P(\mathbf{c}), T(\mathbf{c}), \mathbf{c}) d\mathbf{c} \\ &= \delta_0(\mathbf{r}) + f(\lambda) \times \int_{C(\mathbf{r})} g(P(\mathbf{c}), T(\mathbf{c}), \mathbf{c}) d\mathbf{c},\end{aligned}\quad (1.42)$$

où $\delta_0(\mathbf{r})$ est le chemin optique qu'aurait emprunté la lumière s'il n'y avait pas de matière dans l'atmosphère. En considérant que l'étoile est située à l'infini alors $\delta_0(\mathbf{r}) = \text{cste}$ sur la pupille. Cette constante peut-être considérée comme nulle, $\delta_0(\mathbf{r}) = 0$, puisque l'image est insensible aux constantes globales de la phase (piston). En pratique même si l'étoile n'est pas située à l'infini, les variations de $\delta_0(\mathbf{r})$ sont négligeables. On déduit donc finalement de l'équation (1.42) la relation chromatique différentielle du chemin optique :

$$\begin{aligned}\delta(\lambda, \mathbf{r}) &= f(\lambda) \times \frac{[\delta(\lambda_2, \mathbf{r}) - \delta(\lambda_1, \mathbf{r})]}{f(\lambda_2) - f(\lambda_1)} \\ &= (n(\lambda) - 1) \times \frac{[\delta(\lambda_2, \mathbf{r}) - \delta(\lambda_1, \mathbf{r})]}{n(\lambda_2) - n(\lambda_1)},\end{aligned}\quad (1.43)$$

où la valeur d'indice de réfraction relatif $n(\lambda_i) - 1$ peut-être calculée pour une valeur arbitraire de température et de pression. Par exemple $[n(\lambda) - 1]$ est représentée sur la figure 1.8 pour de l'air sec dans les conditions normales de pression et de température, avec [Allen, 1976] :

$$[n(\lambda) - 1] \times 10^6 = 64.318 + \frac{29498.1}{146 - 1/\lambda^2} + \frac{255.4}{41 - 1/\lambda^2}. \quad (1.44)$$

et λ en μm .

1.4.4.2 Mesure chromatique différentielle de la pente de la surface d'onde

Le principe de l'étoile laser polychromatique est d'estimer le déplacement d_λ d'un objet à observer, à partir des déplacements apparents $d_{\lambda_2}^{\text{app}}$ et $d_{\lambda_1}^{\text{app}}$ d'une étoile artificielle émettant de la lumière à plusieurs longueurs d'onde. En notant d_{λ_0} le déplacement subit à l'aller par le faisceau laser, on voit que la différence chromatique de déplacement mesurée sur l'étoile laser polychromatique est la même que celle de l'objet :

$$\begin{aligned}d_{\lambda_2}^{\text{app}} - d_{\lambda_1}^{\text{app}} &= (d_{\lambda_2} + d_{\lambda_0}) - (d_{\lambda_1} + d_{\lambda_0}) \\ &= d_{\lambda_2} - d_{\lambda_1}.\end{aligned}\quad (1.45)$$

Le déplacement d_λ est ensuite obtenu grâce à la relation différentielle chromatique du chemin optique. En effet comme la relation (1.43) fait intervenir les trois chemins

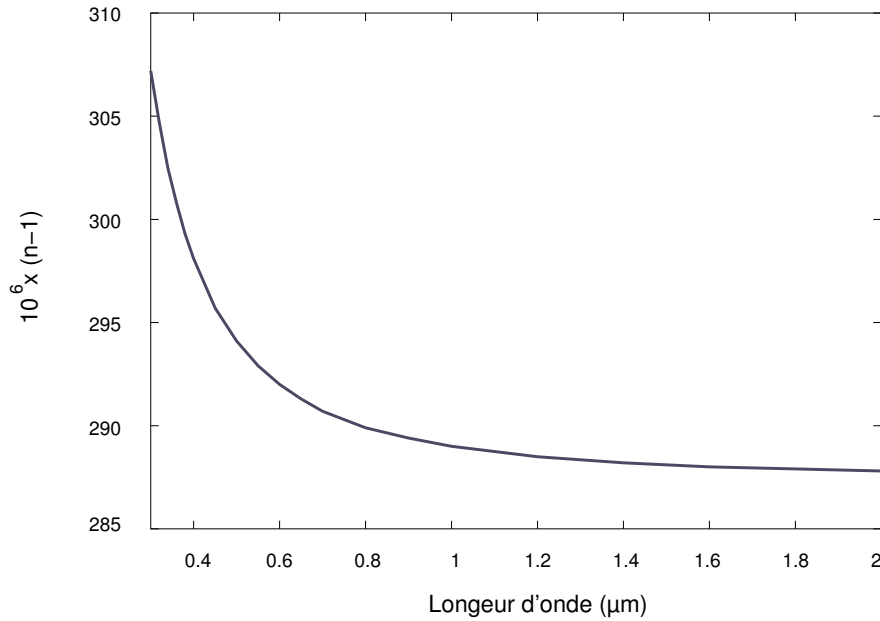


FIG. 1.8 – Variation chromatique de l'indice de réfraction de l'air pour les conditions normales de température et de pression : $P = 1.013 \times 10^5$ Pa et $T = 15^\circ\text{C}$.

optiques chromatiques de façon linéaire, elle peut également être utilisée pour établir le chromatisme de grandeurs définies linéairement à partir du chemin optique, notamment la pente et l'angle d'arrivée de la surface d'onde.

Les coefficients $(\theta_1, \theta_2) \triangleq \theta$ de pente de la surface d'onde, exprimés en unité angulaire de déplacement de l'image à l'aide de la fréquence spatiale $(u_1, u_2) \triangleq \mathbf{u}$ telle que $\mathbf{u} = \mathbf{r}/\lambda$, sont reliés au chemin optique par la relation linéaire :

$$\theta_k(\lambda) = \frac{\iint_{\mathcal{P}} u_k(\mathbf{u}) \times \delta_\lambda(\mathbf{u}) d\mathbf{u}}{\iint_{\mathcal{P}} u_k^2(\mathbf{u}) d\mathbf{u}}, \quad k = 1, 2 \quad (1.46)$$

où \mathcal{P} représente le support de la pupille. En combinant la relation précédente avec l'équation (1.43), on obtient :

$$\theta_k(\lambda) = \frac{n(\lambda) - 1}{n(\lambda_2) - n(\lambda_1)} \times [\theta_k(\lambda_2) - \theta_k(\lambda_1)] . \quad (1.47)$$

Cette relation chromatique différentielle permet d'obtenir une estimation de la pente du front d'onde atmosphérique $\theta_k(\lambda)$ à partir des mesures de pente $\theta_k^{\text{app}}(\lambda_2)$ et $\theta_k^{\text{app}}(\lambda_1)$.

La mesure de déplacement peut être également effectuée avec les angles d'arrivée (α_1, α_2) . Ces angles, exprimés en unité angulaire de déplacement de l'image, sont en effet reliés linéairement au chemin optique par :

$$\alpha_k(\lambda) = \frac{1}{S_{\mathcal{P}}} \times \iint_{\mathcal{P}} \frac{\partial \delta_\lambda(\mathbf{u})}{\partial u_k(\mathbf{u})} d\mathbf{u}, \quad k = 1, 2 \quad (1.48)$$

où $S_{\mathcal{P}}$ est l'aire de la pupille du télescope. A l'analogie de la pente, on peut donc estimer l'angle d'arrivée de la surface d'onde $\alpha_k(\lambda)$, à partir des angles d'arrivée apparents $\alpha_k^{\text{app}}(\lambda_2)$ et $\alpha_k^{\text{app}}(\lambda_1)$ mesurés sur l'étoile laser polychromatique :

$$\alpha_k(\lambda) = \frac{n(\lambda) - 1}{n(\lambda_2) - n(\lambda_1)} \times [\alpha_k^{\text{app}}(\lambda_2) - \alpha_k^{\text{app}}(\lambda_1)] . \quad (1.49)$$

L'angle d'arrivée de la surface d'onde $(\alpha_1, \alpha_2) \triangleq \alpha$ peut-être mesuré directement dans les images, puisqu'il est égal au barycentre de l'image (b_1, b_2) dans le cas où l'amplitude de l'onde ne fluctue pas [Tatarskii, 1971]. En effet en notant $i(\mathbf{x})$ la répartition spatiale d'énergie dans l'image et $I(\mathbf{u})$ sa transformée de Fourier, il vient :

$$\begin{aligned} b_k(\lambda) &= \frac{\iint x_k(\mathbf{x}) \times i(\mathbf{x}) d\mathbf{x}}{\iint i(\mathbf{x}) d\mathbf{x}} \\ &= -\frac{\lambda}{j 2\pi} \times \left(\frac{\partial I_\lambda(\mathbf{u})}{\partial u_k(\mathbf{u})} / I_\lambda(\mathbf{u}) \right)_{\mathbf{u}=\mathbf{0}} . \end{aligned} \quad k = 1, 2 \quad (1.50)$$

Comme on l'a vu à la section 1.1.1, $I(\mathbf{u})$ est l'autocorrélation de l'amplitude complexe transmise par la pupille, à une constante multiplicative près. En supposant que la fonction pupille $P(\mathbf{u})$ est achromatique, réelle et unitaire, et en supposant que l'amplitude de l'onde EM ne fluctue pas sur la pupille, $I_\lambda(\mathbf{u})$ s'écrit :

$$I_\lambda(\mathbf{u}) = \text{cste} \times \iint_{S_{P \otimes P}} e^{j \frac{2\pi}{\lambda} (\delta_\lambda(\mathbf{u}') - \delta_\lambda(\mathbf{u}' + \mathbf{u}))} d\mathbf{u}' , \quad (1.51)$$

où $S_{P \otimes P}$ est le support de l'autocorrélation de la fonction pupille. En appliquant la dérivée partielle à l'expression précédente il vient :

$$\left(\frac{\partial I_\lambda(\mathbf{u})}{\partial u_k(\mathbf{u})} \right)_{\mathbf{u}=\mathbf{0}} = -\frac{j 2\pi}{\lambda} \times \iint_{\mathcal{P}} \frac{\partial \delta_\lambda(\mathbf{u}')}{\partial u_k(\mathbf{u}')} d\mathbf{u}' . \quad (1.52)$$

Comme $I_\lambda(\mathbf{u}=\mathbf{0})=S_{\mathcal{P}}$, on retrouve l'égalité entre barycentre et angle d'arrivée du chemin optique :

$$\begin{aligned} b_k &= \frac{1}{S_{\mathcal{P}}} \times \iint_{\mathcal{P}} \frac{\partial \delta_\lambda(\mathbf{u})}{\partial u_k(\mathbf{u})} d\mathbf{u} \\ &= \alpha_k \end{aligned} \quad k = 1, 2 . \quad (1.53)$$

1.4.4.3 Cascade radiative de l'étoile laser polychromatique

Tout comme la majorité des étoiles laser monochromatiques, l'étoile polychromatique est créée par excitation laser du sodium mésosphérique. L'excitation se fait cette fois-ci en deux étapes. L'absorption résonnante incohérente de deux photons à 589 puis

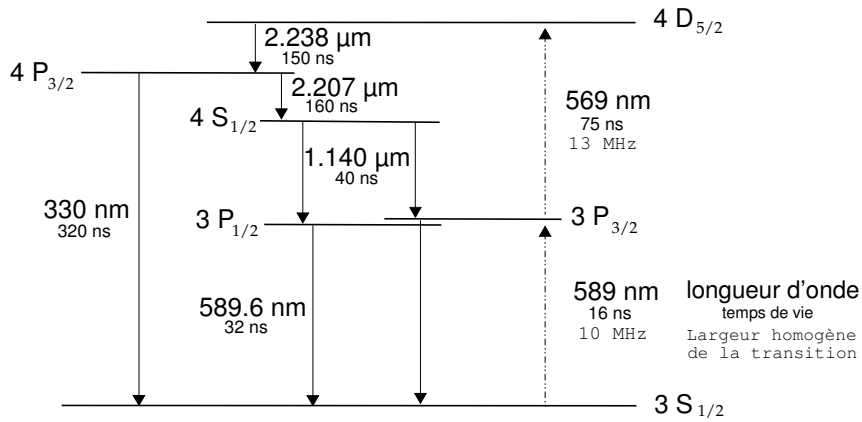


FIG. 1.9 – Cascade radiative polychromatique du sodium

569 nm, permet en effet de porter les électrons de valence du sodium de la configuration $3 S_{1/2}$ vers le niveau $4 D_{5/2}$, à la limite de l'ionisation, et d'engendrer une cascade radiative comprise entre 2.3 et 0.3 μm , comme cela est représenté sur la figure 1.9.

L'indice de réfraction n est que faiblement chromatique, comme on le voit sur la figure 1.8, et il le devient de moins en moins au fur et à mesure que la longueur d'onde augmente. La mesure du déplacement angulaire différentiel gagne donc à être effectuée avec un écart de longueur d'onde le plus grand possible, si les flux monochromatiques sont comparables. En terme d'écart-type, l'erreur finale sur le tip-tilt recherché est multipliée par le facteur chromatique $\gamma_{\lambda_2, \lambda_1}(\lambda) = [n(\lambda) - 1] / [n(\lambda_2) - n(\lambda_1)]$, par rapport à l'erreur commise sur la mesure différentielle, puisque $\theta(\lambda) = \gamma_{\lambda_2, \lambda_1}(\lambda) \times \delta\theta_{\lambda_2, \lambda_1}$. Ce facteur dépend beaucoup du couple (λ_1, λ_2) , et il est calculé dans le tableau ci-dessous pour les différentes longueurs d'onde de la cascade radiative polychromatique du sodium en se basant sur une correction pour $\lambda \sim 500 \text{ nm}$:

$\lambda_1 \setminus \lambda_2$ (nm)	330	589	1 140	2 207
589	25.8			
1 140	19.7	83.3		
2 207	18.5	66.0	319.3	
2 338	18.5	65.5	307.1	8095

On comprend donc tout l'enjeu de trouver des méthodes d'estimations très précises du déplacement des images tavelées, afin de pallier l'inévitable dégradation de la

mesure par le facteur $\gamma_{\lambda_2, \lambda_1}$ calculé dans ce tableau.

En supposant que la mesure du déplacement différentiel est faite très simplement par un barycentre de chaque image, l'écart-type de l'erreur qu'on peut au mieux espérer obtenir sur la mesure différentielle si le bruit de photons domine (cf. borne de Cramér-Rao au chapitre 5) est de l'ordre de :

$$\sigma_{\Delta\theta(\lambda_1, \lambda_2)} = \sqrt{\frac{s_1^2 \times N_{\text{ph}_2} + s_2^2 \times N_{\text{ph}_1}}{N_{\text{ph}_1} N_{\text{ph}_2}}} \quad (1.54)$$

où s_k^2 est la variance de la tache image à la longueur d'onde λ_k et N_{ph_k} le nombre moyen de photons de l'image. La balance entre cette erreur et le facteur $\gamma_{\lambda_2, \lambda_1}(\lambda)$ conduit à considérer les couples (330|589), (330|2338), et (589|2338) nm pour le projet ELP-OA, sachant qu'on peut éventuellement "additionner" les longueurs d'onde 2338 et 2207 μm .

Chapitre 2

Introduction aux problèmes inverses

Je présente dans ce chapitre une formulation probabiliste des problèmes inverses, qui s'appuie en partie sur la théorie établie par Tarantola [1987]. Cette formulation est assez abstraite par rapport au cas concret de reconstruction de phase auquel je m'intéresse dans les chapitres suivants. Mais elle permet à mon sens, en tant que vision théorique et abstraite, de mieux en apprécier les alternatives et améliorations possibles.

Je rappelle également brièvement les principes et caractéristiques des différentes techniques associées à l'optimisation des problèmes inverses non-linéaires, ce qui sert d'assise aux développements particuliers d'optimisation que j'ai faits pour la reconstruction de phase.

2.1 Préambule

2.1.1 Description phénoménologique

En toute généralité, un problème physique *inverse* consiste à retrouver des causes, à partir de l'observation des effets. Les effets, d'un point de vue systémique, sont dits observés en “sortie” d'un ensemble de processus physiques ou “système” ; lequel est supposé isolé, pour que le problème ait un sens. C'est-à-dire que le système contient les principales causes de ses effets observés et n'est donc pas influencé par les événements qui lui sont extérieurs.

Les effets sont plus généralement appelés observables ou données \mathbf{d} , tandis que les causes qu'on recherche sont appelées les paramètres \mathbf{p} . Le terme d'*inverse* désigne donc l'inversion qu'il faut faire de la modélisation directe décrivant les liens de causes à effets :

$$\mathcal{M} : \mathbf{p} \rightarrow \mathbf{d}, \quad (2.1)$$

modélisation qui a été établie pour le problème physique déterministe de prédiction.

Aussi la première difficulté du problème inverse apparaît dans l'absence possible d'équivalence entre l'association des causes et de leurs effets. Le physicien-chimiste

qui ne fait pas de physique quantique, peut en effet s'appuyer sur du déterminisme expérimental - de l'injectivité en termes mathématiques - pour prévoir les résultats de ses expériences, en contrôlant les paramètres effectifs de son expérience. Tandis que s'il recherche à l'*inverse* des paramètres inaccessibles d'une expérience réalisée, et reliés par un modèle à ses observations, il peut trouver plusieurs possibilités de paramètres, parce que le modèle \mathcal{M} n'est pas surjectif. Il peut alors essayer de construire un nouveau modèle reliant de façon univoque les observations à ses paramètres, mais cela est rarement réalisable ou réalisé. Il peut autrement régulariser son problème inverse, c'est-à-dire ajouter des règles que doivent a priori respecter les paramètres qu'il recherche, afin de restreindre les ambiguïtés de "l'inversion" du modèle direct.

2.1.2 Les débuts historiques de formalisation

Les débuts de formalisation pour les problèmes inverses sont fréquemment datés de vers la fin du XVIII^{ème} siècle [Stigler, 1986], avec les travaux de Laplace, Legendre et Gauss dans l'ordre chronologique. On cherchait alors à estimer la forme de la Terre grâce à des données géodésiques et à estimer des trajectoires de planètes et de comètes à partir d'observations lumineuses astronomiques. Ces scientifiques commencèrent pour cela à faire de l'optimisation en ajustant des paramètres \mathbf{p} pour que l'erreur entre le modèle discret paramétré $\mathbf{m}(\mathbf{p})$ et les données \mathbf{d} recueillies soit minimisée. Laplace utilisait une norme absolue pour mesurer l'erreur en minimisant $\sum_i |d_i - m_i(\mathbf{p})|$, tandis que Legendre et Gauss utilisaient une norme quadratique en minimisant $\sum_i (d_i - m_i(\mathbf{p}))^2$. Les méthodes dérivées des moindres carrés de Gauss ont ensuite connu un développement largement majoritaire à cause de leur simplicité de mise en œuvre algébrique.

Le développement de la théorie des probabilités à partir des travaux antérieurs de Fermat, Pascal puis Bernoulli a ensuite permis de réinterpréter ces choix de mesure de l'erreur, en termes d'hypothèses sur la probabilité des données. Tandis que les travaux mathématiques de Thomas Bayes puis de Laplace ont permis d'introduire la notion de probabilité ou mesure a posteriori, et d'ajouter la régularisation des paramètres dans le problème inverse.

2.2 Formalisation probabiliste

A cause de la nature fluctuante des expérimentations et des paramètres physiques, on exprime en terme de densité de probabilité plutôt que par des égalités strictes, l'adéquation $\mathbf{d} \approx \mathbf{m}(\mathbf{p})$ entre les données et le modèle paramétré.

De même pour l'adéquation entre les paramètres reconstruits et leurs a priori. "L'inversion" de modèle consiste dans ce contexte à rechercher le maximum des densités de probabilités ou maximum de vraisemblance. Toutefois pour résoudre effectivement

un problème inverse, il faut effectuer le choix des bons a priori et surtout celui de leur poids μ , ce qui n'est pas réglé par la formulation bayésienne. Cette subtilité est importante comme on le verra au chapitre suivant.

Comme les termes de *vraisemblance* et *maximum de vraisemblance* (Maximum Likelihood) désignent pour certains l'adéquation $\mathbf{d} \approx \mathbf{m}(\mathbf{p})$, tandis que pour d'autres comme Tarantola, ils peuvent également désigner l'adéquation a priori ou l'adéquation a posteriori — qui est la combinaison des deux précédentes —, je choisis par la suite de désigner explicitement les adéquations auxquelles je fais référence en employant les termes, *expérimental*, *a priori* et *a posteriori*, plutôt que d'utiliser le terme *vraisemblance*.

2.2.1 Induction

L'identification des paramètres suppose comme on l'a vu un premier positionnement strictement inductif. Il faut tout d'abord avoir défini le système physique isolé \mathcal{S} , les paramètres \mathbf{p} qu'on souhaite connaître, et le modèle \mathcal{M} reliant ces paramètres aux observations \mathbf{d} :

$$\begin{array}{ccc} \mathcal{M} : & \mathcal{P} & \mapsto \mathcal{D} \\ & \mathbf{p} & \longrightarrow \mathbf{d} \end{array} \quad (2.2)$$

Je considère ici le cas simplifié où l'espace des données \mathcal{D} et l'espace des paramètres \mathcal{P} sont des espaces euclidiens de dimension finie.

Les observations \mathbf{d} correspondent à des mesures, c'est-à-dire qu'on considère que l'information n'est pas recueillie par l'homme directement mais par un outil ou instrument qu'il a conçu pour dépasser les limitations de ses propres sens. On pourrait aussi s'intéresser à la tacite faculté du cerveau humain à résoudre les problèmes inverses en injectant de bons ou de mauvais a priori à partir des informations ressenties, mais ce n'est pas le cadre de l'étude.

Les paramètres \mathbf{p} correspondent parfois à une reparamétrisation de l'ensemble principal des paramètres qu'on souhaite connaître. En effet il est nécessaire de caractériser le système en un nombre minimum de paramètres indépendants pour limiter les ambiguïtés de "l'inversion".

2.2.2 Espaces probabilisés des observables et des paramètres

On introduit une probabilité $\text{Pr}_{\mathcal{D}}$ sur \mathcal{D} traduisant la nature aléatoire des processus physiques du système, que ce soit au niveau des processus de création du signal, de sa transmission ou de sa réception par l'instrument de mesure. Cette probabilité associe par exemple à une réalisation de données \mathbf{d} , le nombre normalisé d'occurrences pour un grand nombre d'observations. Cette probabilité peut aussi contenir une certaine part de subjectivité qui provient de considérations mathématiques ou physiques. De même on introduit une probabilité $\text{Pr}_{\mathcal{P}}$ sur les paramètres \mathbf{p} .

2.2.3 Densité de probabilité expérimentale

Connaissant la modélisation directe $\mathbf{d}=\mathbf{m}(\mathbf{p})$, l'information expérimentale à cause de sa nature fluctuante, s'exprime sur le produit des espaces probabilisés $\mathcal{P} \times \mathcal{D}$ par la densité de probabilité expérimentale :

$$f_{\text{exp}}(\mathbf{d}, \mathbf{p}) = f_{\mathcal{D}|\mathcal{P}}(\mathbf{d}|\mathbf{p}) h_{\mathcal{D}}(\mathbf{p}), \quad (2.3)$$

où $h_{\mathcal{D}}(\mathbf{p})$ est la densité de probabilité marginale homogène et

$$f_{\mathcal{D}|\mathcal{P}}(\mathbf{d}|\mathbf{p}) \triangleq f_{\mathcal{D}|\mathcal{P}}[\mathbf{d} | \mathbf{d} \approx \mathbf{m}(\mathbf{p})] . \quad (2.4)$$

La densité de probabilité homogène est la densité de probabilité qui est proportionnelle au volume. Elle traduit ici l'absence d'information expérimentale sur les paramètres qui soit indépendante des données. Dans toute la suite on considère que les paramètres appartiennent à un espace vectoriel fini et on définit le volume élémentaire dV à partir des vecteurs élémentaires de cet espace vectoriel, de sorte que cette densité est une constante.

2.2.4 Densité de probabilité a priori

Comme précédemment expliqué, la régularisation du problème inverse à l'aide des a priori vise à lever les ambiguïtés de "l'inversion" du modèle direct. Elle sert en fait aussi à améliorer le conditionnement pour l'inversion du modèle direct, c'est-à-dire à empêcher que des faibles variations dans les données qui soient liées au bruit uniquement, correspondent à de trop grandes variations dans les paramètres (voir section 2.5.2)

On suppose que l'information a priori vient compléter l'information expérimentale de façon indépendante, de sorte que les a priori sur les données et les a priori sur les paramètres sont indépendants, ce qui conduit à

$$f_{\text{prior}}(\mathbf{d}, \mathbf{p}) = f_{\text{prior}}(\mathbf{d}) \times f_{\text{prior}}(\mathbf{p}) . \quad (2.5)$$

Lorsqu'on ne sait pas corriger les données, on ne considère que des a priori sur les paramètres. Cette absence d'a priori sur les données se traduit par une densité de probabilité homogène, prise constante par la suite pour le cas simplifié des espaces vectoriels et du volume élémentaire.

2.2.5 Densité de probabilité a posteriori

En supposant l'absence de recoupement entre a priori et expérience, la combinaison de l'information des observations, des a priori et de la modélisation théorique s'exprime par la densité de probabilité jointe a posteriori

$$f_{\text{post}}(\mathbf{d}, \mathbf{p}) = \text{cste} \times f_{\text{exp}}(\mathbf{d}, \mathbf{p}) \times f_{\text{prior}}(\mathbf{d}, \mathbf{p}) . \quad (2.6)$$

L'information a posteriori dans l'espace des paramètres s'exprime donc finalement par

$$f_{\text{post}}(\mathbf{p}) = \text{cste} \times f_{\text{prior}}(\mathbf{p}) \times \int_{\mathcal{D}} f_{\text{exp}}(\mathbf{d}|\mathbf{p}) \, d\mathbf{d}. \quad (2.7)$$

On note couramment $L(\mathbf{p}) = \int_{\mathcal{D}} f_{\text{exp}}(\mathbf{d}|\mathbf{p}) \, d\mathbf{d}$ la fonction mesurant la vraisemblance (Likelihood) expérimentale entre les données et les paramètres. Pour les raisons détaillées au paragraphe 2.2, je désigne par la suite celle-ci par la notation plus explicite $f_{\text{exp}}(\mathbf{p}) \triangleq L(\mathbf{p})$.

2.3 Remarques

2.3.1 Choix du maximum

Le choix de la plus forte adéquation a posteriori du modèle paramétré avec les données (maximum de $f_{\text{post}}(\mathbf{p})$), est le choix le plus intuitif pour trouver la meilleure estimation \mathbf{p}^+ des paramètres. Toutefois on imagine que ce choix ne soit pas forcément le bon dans certains cas et qu'il faille par exemple lui préférer le maximum du deuxième mode ou une valeur intermédiaire.

La modélisation de l'adéquation a posteriori sous la forme de densité de probabilité permet en théorie de calculer jusqu'à la probabilité $P_{\mathbf{p}}(A) = \int_A f_{\text{post}}(\mathbf{p}) \, d\mathbf{p}$ que les paramètres appartiennent à un sous-ensemble $A \subset \mathcal{P}$. Dans le cas où l'on effectue un travail de caractérisation avec des données simulées et où les paramètres réels du problème inverse sont en fait connus du simulateur, on peut penser mener une étude approfondie de la densité de probabilité a posteriori et déterminer le point de *meilleure* adéquation au sens de celle qui permet d'estimer les paramètres le plus précisément. Toutefois ce genre d'étude n'est envisageable que lorsque le nombre de paramètres est restreint et que le calcul point à point de f_{post} est rapide. Ce n'est pas le cas du problème de reconstruction de phase auquel je me suis attaqué dans les chapitres suivants. J'ai donc a priori supposé que le maximum était un bon critère, et vérifié a posteriori la validité de ce choix.

2.3.2 Importance du volume choisi

Lorsque le volume est défini dans l'espace des paramètres de façon non élémentaire, il convient de définir le maximum de vraisemblance comme le maximum de la quantité invariante associée $f(\mathbf{p})/h(\mathbf{p})$ appelée densité de probabilité volumétrique, qui est uniquement définie d'après le théorème de Radon-Nikodym et ne dépend plus du volume choisi. Sans cette normalisation par la densité de probabilité homogène, le maximum peut en effet correspondre à des paramètres forts différents.

2.3.3 Fonction d'adéquation a posteriori

Plutôt que de chercher le maximum de la densité de probabilité f_{post} , on cherche souvent de façon équivalente le maximum de son image par une fonction bijective monotone, soit parce que celle-ci conduit à des calculs numériques plus simples dans le cas de l'optimisation non-linéaire (sans débordement numérique), soit parce que le calcul analytique du maximum est simplifié en considérant cette fonction dans le cas linéaire. Par exemple pour des incertitudes gaussiennes sur les adéquations, on minimise

$$f'_{\text{post}}(\mathbf{p}) = -\ln(f_{\text{post}}(\mathbf{p})), \quad (2.8)$$

plutôt que de maximiser $f_{\text{post}}(\mathbf{p})$.

2.3.4 Poids de la régularisation

Pour améliorer le conditionnement local du modèle expérimental lors de l'optimisation de $f'_{\text{post}}(\mathbf{p})$, et pour favoriser le choix de certains paramètres, lorsqu'il y a des ambiguïtés par rapport à l'adéquation expérimentale, on choisit souvent de renforcer le poids de l'adéquation a priori dans l'équation (2.7).

En notant μ l'hyperparamètre qui règle le poids de la régularisation, on cherche ainsi plutôt à minimiser

$$f'_{\text{post}}(\mathbf{p}) = f'_{\text{exp}}(\mathbf{p}) + \mu \times f'_{\text{prior}}(\mathbf{p}). \quad (2.9)$$

2.4 Cas particulier des fluctuations gaussiennes

Dans le cas où les incertitudes de la modélisation et de l'expérience suivent des statistiques gaussiennes de matrice de covariance respectives \mathbf{C}_m et \mathbf{C}_{exp} , on montre [Tarantola, 1987] que les opérateurs de covariance se combinent. La densité de probabilité a posteriori s'écrit alors

$$f_{\text{post}}(\mathbf{p}) = \text{cste} \times \exp \left[-\frac{1}{2} (\mathbf{d} - \mathbf{m}(\mathbf{p}))^\top \cdot \mathbf{C}_n^{-1} \cdot (\mathbf{d} - \mathbf{m}(\mathbf{p})) \right] \times f_{\text{prior}}(\mathbf{p}), \quad (2.10)$$

où $\mathbf{C}_n = \mathbf{C}_m + \mathbf{C}_{\text{exp}}$. Lorsque l'information a priori est également gaussienne on obtient

$$f'_{\text{post}}(\mathbf{p}) \propto (\mathbf{d} - \mathbf{m}(\mathbf{p}))^\top \cdot \mathbf{C}_n^{-1} \cdot (\mathbf{d} - \mathbf{m}(\mathbf{p})) + (\mathbf{p}_{\text{prior}} - \mathbf{p})^\top \cdot \mathbf{C}_{\text{prior}}^{-1}(\mu) \cdot (\mathbf{p}_{\text{prior}} - \mathbf{p}). \quad (2.11)$$

De plus, lorsque la relation entre le modèle et les données est linéaire :

$$\mathbf{m}(\mathbf{p}) = \mathbf{M} \cdot \mathbf{p}, \quad (2.12)$$

on obtient analytiquement l'expression du maximum a posteriori :

$$\tilde{\mathbf{p}}_{\text{post}} = \mathbf{p}_{\text{prior}} + \left(\mathbf{M}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{M} + \mathbf{C}_{\text{prior}}^{-1} \right)^{-1} \cdot \mathbf{M}^\top \cdot \mathbf{C}_n^{-1} \cdot (\mathbf{d} - \mathbf{M} \cdot \mathbf{p}_{\text{prior}}), \quad (2.13)$$

et celle de la covariance a posteriori :

$$\mathbf{C}_{\mathbf{p}_{\text{post}}} = \left(\mathbf{M}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{M} + \mathbf{C}_{\text{prior}}^{-1} \right)^{-1}. \quad (2.14)$$

On voit dans cette dernière expression que $\mathbf{C}_{\mathbf{p}_{\text{post}}} \leq \mathbf{C}_{\mathbf{p}_{\text{prior}}}$ traduit le gain apporté par l'information expérimentale.

2.5 Optimisation pseudo-continue pour les problèmes inverses non-linéaires

Lorsque le modèle et les paramètres ne sont pas reliés linéairement et/ou lorsque les fluctuations ne sont pas gaussiennes, il n'existe en général pas de relation analytique permettant d'obtenir les paramètres \mathbf{p}^+ qui maximisent l'adéquation f_{post} . Ces paramètres sont donc obtenus par optimisation itérative, c'est-à-dire en sélectionnant des paramètres qui améliorent $f_{\text{post}}(\mathbf{p})$ progressivement. Le choix des paramètres d'une itération à l'autre, peut se faire soit à l'aide d'un développement local de Taylor de $f_{\text{post}}(\mathbf{p})$, en anticipant par continuité les variations de l'adéquation avec les variations des paramètres, soit de façon discontinue avec des méthodes pseudo-aléatoires de type Monte-Carlo, lorsque la continuité de l'adéquation est vraiment trop faible.

Au-delà des ambiguïtés strictes de l'adéquation, l'optimisation d'un problème non-linéaire par une méthode pseudo-continue, peut être compliquée par la possibilité de convergence vers un extremum local de l'adéquation, lorsque celle-ci est multimodale. Le problème de reconstruction de phase que je traite par la suite est notamment un bon exemple d'un tel cas. En plus de vérifier que le maximum global est un bon critère pour estimer les paramètres recherchés, il faut alors trouver un moyen de l'atteindre en évitant les extrema secondaires.

2.5.1 Formalisme

Comme on l'a remarqué précédemment, il est plus facile de chercher le minimum du logarithme négatif de f_{post} , que de chercher directement son maximum. En convention avec les techniques d'optimisation pour les problèmes inverses on parlera donc dans toute la suite de la minimisation de $f' = -\ln(f)$ plutôt que de la maximisation de f . Pour alléger la notation j'omets également le prime de f dans toute cette section, bien que je fasse référence à f' telle que je viens de la définir. Cette convention est d'ailleurs plus en accord avec les premières formulations des problèmes inverses, en terme de minimisation d'une erreur entre les données et le modèle paramétré, à l'aide d'une fonction de "coût" ou fonction de "pénalisation".

Les méthodes d'optimisation pseudo-continues sont comme on l'a vu vouées à trouver un minimum de f , en exploitant itérativement la continuité locale de $f(\mathbf{p})$ à

l'aide d'un développement de Taylor. On utilise en général un développement à l'ordre 2 au maximum :

$$f(\mathbf{p} + \delta\mathbf{p}) = f(\mathbf{p}) + \mathbf{g}^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H} \cdot \delta\mathbf{p} + o(\|\delta\mathbf{p}\|^2), \quad (2.15)$$

où $\mathbf{g} \triangleq \nabla_{\mathbf{p}} f$ et $\mathbf{H} \triangleq \nabla_{\mathbf{p}}^2 f$ sont le gradient et la hessienne de f (les dérivées partielles d'ordre 1 et 2 de f).

Comme un pas $\delta\mathbf{p}$ minimisant $f(\mathbf{p} + \delta\mathbf{p})$, minimise également

$$\delta f(\mathbf{p}) = f(\mathbf{p} + \delta\mathbf{p}) - f(\mathbf{p}), \quad (2.16)$$

on cherche plus couramment à minimiser cette dernière fonction qui ne contient pas de constante inutile. L'approximation quadratique de f qui sert au final à la minimisation s'écrit donc

$$\delta f^{\text{quad}}(\mathbf{p}, \delta\mathbf{p}) = \mathbf{g}^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H} \cdot \delta\mathbf{p}. \quad (2.17)$$

2.5.2 Algorithme de Newton

En toute généralité on montre que pour qu'un point soit le minimum local d'une fonction de plusieurs variables, il est suffisant que le gradient s'annule en ce point et que la matrice hessienne y soit positive définie [Gockenbach, 2005]. En considérant le cas simple d'une fonction convexe, l'algorithme de Newton consiste ainsi en une suite de pas $\delta\mathbf{p}$ vérifiant

$$\nabla_{\delta\mathbf{p}} \delta f^{\text{quad}}(\mathbf{p}, \delta\mathbf{p}) = 0. \quad (2.18)$$

En effet, comme f est convexe on a par ailleurs $\nabla_{\delta\mathbf{p}}^2 \delta f^{\text{quad}}(\mathbf{p}, \delta\mathbf{p}) = \mathbf{H}$ positive définie. En développant l'équation (2.18) on obtient le pas de Newton servant de mise à jour itérative des paramètres :

$$\delta\mathbf{p}^{\text{Newton}} = -\mathbf{H}^{-1}(\mathbf{p}) \cdot \mathbf{g}(\mathbf{p}). \quad (2.19)$$

La force de l'algorithme de Newton tient à ses propriétés de convergence. En effet on démontre la convergence quadratique de l'algorithme vers un minimum local de f , si f est Lipschitz-continue, i.e. si $\exists C \in \mathbb{R}^+$ tel que $|f(\mathbf{x}) - f(\mathbf{y})| < C\|\mathbf{x} - \mathbf{y}\| \quad \forall(\mathbf{x}, \mathbf{y})$, ce qui est le cas des fonctions dont la dérivée première est bornée notamment et a fortiori des fonctions deux fois continûment dérivables.

En pratique deux problèmes limitent malheureusement l'intérêt de la méthode de Newton. Tout d'abord le calcul de la hessienne \mathbf{H} est un calcul lourd qui doit être effectué à chaque itération. Cela devient vite impossible pour un grand nombre de paramètres, soit parce que la matrice est trop longue à calculer, soit tout simplement parce que la quantité de mémoire nécessaire pour la stocker est trop importante (en effet on a $\mathbf{H} \in \mathbb{R}^{N \times N}$, où N est le nombre de paramètres). Le deuxième problème est

que la résolution numérique de l'équation devient instable lorsque la matrice \mathbf{H} a un mauvais (élevé) conditionnement

$$\text{cond}(\mathbf{H}) = \|\mathbf{H}^{-1}\| \times \|\mathbf{H}\|, \quad (2.20)$$

lequel se calcule pour une norme quadratique (norme L_2), comme le rapport des valeurs propres extrêmes de \mathbf{H} , ou comme le rapport des valeurs singulières extrêmes pour une norme de Laplace (norme L_1). La résolution numérique peut même devenir impossible lorsque \mathbf{H} n'est pas positive définie. On ne peut en effet alors plus calculer l'inverse de \mathbf{H} pour le pas de Newton.

2.5.3 Algorithmes de quasi-Newton

Les algorithmes de minimisation de quasi-Newton effectuent une suite de pas de type Newton vers un minimiseur de f , en utilisant des approximations $\mathbf{A}(\mathbf{p})$ de $\mathbf{H}(\mathbf{p})$ qui sont positives-définies et bien conditionnées. Cela assure en effet la convergence vers un minimum local de f [Nocedal, 1992]. L'approximation quadratique de Taylor servant aux algorithmes de quasi-Newton se réécrit

$$\delta f^{\text{quad}}(\mathbf{p}, \delta \mathbf{p}) = \mathbf{g}^\top \cdot \delta \mathbf{p} + \frac{1}{2} \delta \mathbf{p}^\top \cdot \mathbf{A} \cdot \delta \mathbf{p}, \quad (2.21)$$

à partir de laquelle on déduit le pas de quasi-Newton de mise à jour des paramètres :

$$\delta \mathbf{p}^{\text{QN}} = -\mathbf{A}^{-1}(\mathbf{p}) \cdot \mathbf{g}(\mathbf{p}). \quad (2.22)$$

On a également parfois recours à des préconditionneurs qui permettent d'améliorer le conditionnement de \mathbf{A} , une fois que celle-ci a été bien choisie pour approximer \mathbf{H} .

L'algorithme de plus grande pente est par exemple l'algorithme de quasi-Newton le plus ancien et le plus simple. Il consiste à approximer la matrice hessienne par la matrice identité :

$$\mathbf{B} \triangleq \mathbf{A}^{-1} \propto \mathbf{I}. \quad (2.23)$$

Les pas de mise à jour des paramètres s'écrivent alors

$$\delta \mathbf{p}^{\text{pgp}} = -\beta(\mathbf{p}) \times \mathbf{g}(\mathbf{p}) \quad (2.24)$$

où $\beta(\mathbf{p})$ est un scalaire réglant la longueur du pas. L'approximation de la hessienne par $\beta \mathbf{I}$ est cependant trop brutale et ne permet pas une convergence rapide vers un minimum local pour de nombreux problèmes d'optimisation.

2.5.3.1 Algorithme de Levenberg-Marquardt

L'approximation de Levenberg-Marquardt, parfois aussi appelée approximation de Gauss-Newton, est l'approximation \mathbf{A} de \mathbf{H} de référence pour minimiser des fonctions quadratiques de type

$$f^{\text{quad}}(\mathbf{p}) = \frac{1}{2} \mathbf{e}(\mathbf{p})^\top \cdot \mathbf{W} \cdot \mathbf{e}(\mathbf{p}), \quad (2.25)$$

où \mathbf{W} est symétrique et représente en général la matrice de covariance inverse des fluctuations. Le plus souvent \mathbf{e} est un vecteur *résidu*, et f^{quad} est appelée la fonction des moindres carrés de \mathbf{e} . Par exemple

$$\mathbf{e}_{\text{ML}} = \mathbf{d} - \mathbf{m}(\mathbf{p}) \quad (2.26)$$

$$\mathbf{e}_{\text{prior}} = \mathbf{p} - \mathbf{p}_{\text{prior}}, \quad (2.27)$$

pour l'optimisation des adéquations expérimentales et a priori sous l'hypothèse des fluctuations gaussiennes. Lorsque \mathbf{e} est un résidu, f^{quad} est alors une fonction de moindres carrés.

La matrice hessienne de f^{quad} s'écrit

$$\mathbf{H} = \mathbf{J}^\top \cdot \mathbf{W} \cdot \mathbf{J} + \mathbf{S}, \quad (2.28)$$

où $[\mathbf{J}]_{i,j} = \frac{\partial e_i}{\partial p_j}$ est la matrice jacobienne et

$$[\mathbf{S}]_{i,j} = \sum_{k,l} e_k(\mathbf{p}) \times W_{k,l} \times \frac{\partial^2 e_l(\mathbf{p})}{\partial p_i \partial p_j}. \quad (2.29)$$

L'approximation de Levenberg-Marquardt consiste à ne garder que le premier terme de l'équation (2.28) comme approximation de \mathbf{H} :

$$\mathbf{A} = \mathbf{J}^\top \cdot \mathbf{W} \cdot \mathbf{J}, \quad (2.30)$$

qui est positive semi-définie par construction, et permet d'éviter le calcul souvent fastidieux de $\nabla_{\mathbf{p}}^2 \mathbf{e}(\mathbf{p})$. De plus cette approximation a le double avantage de tirer rapidement vers le minimum lorsque \mathbf{p} est loin de celui-ci, tout en constituant une très bonne approximation de \mathbf{H} lorsque \mathbf{p} est proche du minimum [Nocedal and Wright, 2006]. Malheureusement l'approximation de Levenberg-Marquardt est parfois trop coûteuse en stockage voire en temps de calcul pour les problèmes à grand nombre de degrés de liberté.

2.5.3.2 Algorithme de quasi-Newton partitionné

Pour les problèmes à très grand nombre de degrés de liberté, on peut utiliser une reformulation creuse de la méthode de quasi-Newton, lorsque la matrice hessienne

est creuse. On montre en effet dans ce cas-là [Griewank and Toint, 1981] que f est partiellement séparable :

$$f(\mathbf{p}) = \sum_{i=1}^{ne} f_i(\mathbf{p}), \quad (2.31)$$

où chaque fonction ne dépend que d'un nombre restreint de paramètres. La mise à jour se fait alors [Toint, 1983, 1986] en fonction des hessiennes \mathbf{H}^i (ou de leurs approximations creuses \mathbf{A}^i) de chaque f_i par

$$\left(\sum_i^{ne} \mathbf{H}^i(\mathbf{p}) \right) \cdot \delta(\mathbf{p}) = -\mathbf{g}_k. \quad (2.32)$$

2.5.3.3 Algorithmes de métrique variable.

Pour approximer la hessienne asymptotiquement et obtenir la convergence super-linéaire dans des problèmes à grand nombre de degrés de liberté, Davidon [1959] puis Fletcher and Powell [1963] ont proposé de mettre à jour l'approximation \mathbf{A} de la hessienne, en collectant l'information sur la courbure au fur et à mesure des itérations accomplies. En effet les variations du gradient sont reliées à la hessienne par

$$\mathbf{g}^{(k+1)} - \mathbf{g}^{(k)} = \mathbf{H}^{(k)} \cdot (\mathbf{p}^{(k+1)} - \mathbf{p}^{(k)}) + o(\|\mathbf{p}^{(k+1)} - \mathbf{p}^{(k)}\|), \quad (2.33)$$

ce qui conduit directement à mettre à jour l'approximation de la Hessienne à l'itération $(k+1)$ par une approximation qui satisfait

$$\mathbf{y}^{(k)} = \mathbf{A}^{(k+1)} \cdot \mathbf{s}^{(k)}, \quad (2.34)$$

où $\mathbf{y}^{(k)} = \mathbf{g}^{(k+1)} - \mathbf{g}^{(k)}$ et $\mathbf{s}^{(k)} = \mathbf{p}^{(k+1)} - \mathbf{p}^{(k)}$. La condition (2.34) est plus connue sous le nom d'équation de la sécante. Cette condition sert de base à l'ensemble des méthodes de métrique variable. Généralement on ajoute à cette condition que $\mathbf{A}^{(k+1)}$ soit symétrique et que $\mathbf{A}^{(k+1)} - \mathbf{A}^{(k)}$ soit de rang faible.

BFGS est l'algorithme de métrique variable le plus populaire. Son nom vient des quatre chercheurs Broyden, Fletcher, Goldfarb et Schano, qui l'ont proposé à peu près en même temps. La mise à jour de la hessienne à l'itération $(k+1)$ se fait par

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} - \frac{\mathbf{A}^{(k)} \cdot \mathbf{s}^{(k)} \cdot \mathbf{s}^{(k)\top} \cdot \mathbf{A}^{(k)}}{\mathbf{s}^{(k)\top} \cdot \mathbf{A}^{(k)} \cdot \mathbf{s}^{(k)}} + \frac{\mathbf{y}^{(k)} \cdot \mathbf{y}^{(k)\top}}{\mathbf{y}^{(k)\top} \cdot \mathbf{s}^{(k)}}. \quad (2.35)$$

Pour les problèmes à grand nombre de degrés de liberté, on choisit ainsi de ne stocker que les quelques derniers vecteurs $(\mathbf{y}^{(i)}, \mathbf{s}^{(i)})$ (de l'ordre de ~ 5), au lieu d'une hessienne $\mathbf{A}^{(k)}$ beaucoup plus volumineuse, pour pouvoir calculer le pas de quasi-Newton (et on se contente d'une approximation $\mathbf{A}^{(0)}$ diagonale).

L'algorithme alternatif de BFGS est l'algorithme de mise à jour symétrique de rang 1, proposé par Davidon [1959] (la mise à jour de BFGS est symétrique de rang 2). La mise à jour de la hessienne à l'itération $(k+1)$ se fait par

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \frac{(\mathbf{y}^{(k)} - \mathbf{A}^{(k)} \cdot \mathbf{s}^{(k)}) \cdot (\mathbf{y}^{(k)} - \mathbf{A}^{(k)} \cdot \mathbf{s}^{(k)})^\top}{(\mathbf{y}^{(k)} - \mathbf{A}^{(k)} \cdot \mathbf{s}^{(k)})^\top \cdot \mathbf{s}^{(k)}}. \quad (2.36)$$

2.5.4 Algorithme des gradients conjugués

L'introduction des gradients conjugués non-linéaires en 1960 par Fletcher et Reeves a eu un impact considérable sur l'essor de l'optimisation non-linéaire à grand nombre de paramètres, offrant une alternative à convergence beaucoup plus rapide que la méthode de plus grande pente, et à moindre coût de calcul et de stockage par rapport aux algorithmes de quasi-Newton de type Levenberg-Marquardt. La méthode est en fait dérivée de l'algorithme des gradients conjugués linéaires, proposés une dizaine d'années auparavant pour résoudre rapidement les équations linéaires du type $\mathbf{Ax} = \mathbf{b}$ dans le cas d'un très grand nombre de paramètres. La notion de conjugaison vient de l'algèbre linéaire : deux vecteurs $\mathbf{x}^{(k+1)}$ et $\mathbf{x}^{(k)}$ de E sont en effet dits \mathbf{A} -conjugués (ou \mathbf{A} -orthogonaux) lorsque $\mathbf{x}^{(k+1)\top} \cdot \mathbf{A} \cdot \mathbf{x}^{(k)} = 0$.

L'algorithme des gradients conjugués non-linéaires consiste à enchaîner des pas \mathbf{A} -orthogonaux, en combinant à l'aide d'un poids γ la direction de plus grande pente et celle du pas précédent :

$$\delta \mathbf{p}^{(k+1)} \propto -\mathbf{g}^{(k+1)} + \gamma^{(k+1)} \times \delta \mathbf{p}^{(k)}. \quad (2.37)$$

Les poids utilisés dans l'algorithme de Fletcher et Reeves s'écrivent

$$\gamma_{\text{FR}}^{(k)} = 0 \quad \text{pour } k = 1 \quad (2.38)$$

$$\gamma_{\text{FR}}^{(k)} = \frac{\|\mathbf{g}^{(k)}\|^2}{\|\mathbf{g}^{(k-1)}\|^2} \quad \text{pour } k > 1. \quad (2.39)$$

Polak and Ribière [1969] ont par la suite proposé la pondération alternative

$$\gamma_{\text{PR}}^{(k)} = \frac{\mathbf{g}^{(k)\top} \cdot (\mathbf{g}^{(k)} - \mathbf{g}^{(k-1)})}{\|\mathbf{g}^{(k-1)}\|^2}, \quad (2.40)$$

qui s'avère souvent plus efficace, notamment par son meilleur comportement loin de la solution, et qui est également mieux adaptée que celle de Fletcher et Reeves pour les problèmes non-quadratiques. De nombreuses études des propriétés des gradients conjugués [Gilbert and Nocedal, 1990], ont également permis de prouver qu'un redémarrage périodique de l'algorithme avec la direction de plus grande pente ($\gamma^{(k)}=0$)

assure de meilleures propriétés de convergence, ainsi que de prendre $\gamma^{(k)} = \max(\gamma_{\text{PR}}^{(k)}, 0)$. L'algorithme demeure quoiqu'il en soit moins efficace que les algorithmes de quasi-Newton, lorsque le nombre de paramètres et de données n'est pas trop important pour pouvoir mettre en œuvre ces méthodes d'optimisation.

2.5.5 Ajustement de la longueur de pas

On distingue comme on vient de le voir deux grandes familles d'algorithmes de descente : les algorithmes de quasi-Newton qui utilisent des directions successives de minimisation

$$\mathbf{q}^{(k)} = -\mathbf{A}^{-1} \mathbf{g}^{(k)}, \quad (2.41)$$

où \mathbf{A} est une approximation positive définie de la hessienne ; et les algorithmes de type gradients conjugués pour lesquels

$$\mathbf{q}^{(k)} = -\mathbf{g}^{(k)} + \gamma^{(k)} \mathbf{q}^{(k-1)}. \quad (2.42)$$

Sans longueur de pas appropriée, ces algorithmes ne sont pas tous globalement convergents, c'est-à-dire que les propriétés de convergence vers un minimum local dépendent de l'initialisation des paramètres, ce qui est le cas en particulier de l'algorithme des gradients conjugués. Pour les algorithmes de quasi-Newton on montre qu'il n'y a a priori pas besoin de longueur de pas. En effet, si on cherche après avoir déterminé \mathbf{q}^{QN} , une longueur de pas β^+ , qui minimise δf^{quad} :

$$\beta^+ = \min_{\beta} \left\{ \beta \times \mathbf{g}^{\top} \cdot \mathbf{q}^{\text{QN}} + \frac{\beta^2}{2} \times \mathbf{q}^{\text{QN}\top} \cdot \mathbf{A} \cdot \mathbf{q}^{\text{QN}} \right\}, \quad (2.43)$$

on obtient

$$\beta^+ = -\frac{\mathbf{g}^{\top} \cdot \mathbf{q}^{\text{QN}}}{\mathbf{q}^{\text{QN}\top} \cdot \mathbf{A} \cdot \mathbf{q}^{\text{QN}}} = 1.$$

Cependant l'ajustement de la longueur de pas peut permettre d'accélérer la convergence selon la rugosité locale de la fonction à minimiser, et selon la validité des approximations quadratiques par exemple. La majorité des algorithmes de minimisation utilisent donc une mise à jour des paramètres à l'itération (k) à l'aide d'une longueur de pas $\beta^{(k)}$:

$$\delta \mathbf{p}^{(k)} = \beta^{(k)} \times \mathbf{q}^{(k)}. \quad (2.44)$$

2.5.5.1 Conditions de Wolfe

On choisit souvent une longueur du pas $\beta^{(k)}$ en fonction des deux conditions de Wolfe [1969, 1971] :

$$\begin{aligned} f(\mathbf{p}^{(k)} + \beta^{(k)} \mathbf{q}^{(k)}) &\leq f(\mathbf{p}^{(k)}) + \sigma_1 \beta^{(k)} \mathbf{g}^{(k)\top} \cdot \mathbf{q}^{(k)} \\ \mathbf{g}(\mathbf{p}^{(k)} + \beta^{(k)} \mathbf{q}^{(k)})^{\top} \cdot \mathbf{q}^{(k)} &\geq \sigma_2 \mathbf{g}^{(k)\top} \cdot \mathbf{q}^{(k)}, \end{aligned} \quad (2.45)$$

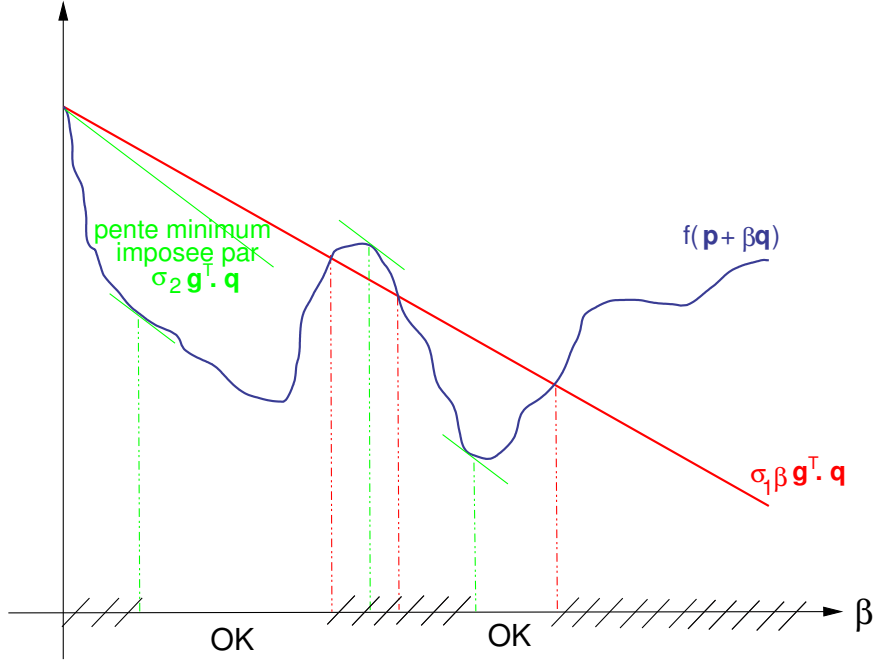


FIG. 2.1 – Représentation graphique des deux conditions de Wolfe

avec $0 < \sigma_1 < \sigma_2 < 1$. En effet celles-ci assurent respectivement une réduction suffisante de f et une longueur de pas suffisante (cf figure 2.1). Lorsque les conditions de Wolfe sont satisfaites, on montre que les algorithmes de plus grande pente, de Levenberg-Marquardt et de métrique variable convergent globalement. Pour les gradients conjugués ce résultat est obtenu à condition d'une réinitialisation périodique avec $\mathbf{q}^{(k)} = -\mathbf{g}^{(k)}$ (i.e. $\gamma^{(k)}=0$).

2.5.5.2 Région de confiance

Une autre façon d'ajuster la longueur du pas est de se baser sur la validité des approximations quadratiques en utilisant la quantité

$$\rho_{\mathbf{p}}(\delta\mathbf{p}) = \frac{\delta f_{\mathbf{p}}(\delta\mathbf{p})}{\delta f_{\mathbf{p}}^{\text{quad}}(\delta\mathbf{p})}, \quad (2.46)$$

qui exprime le rapport de la réduction δf de la fonction à minimiser avec son approximation quadratique δf^{quad} . En effet plus on s'éloigne du point \mathbf{p} plus il y a de chances que les approximations quadratiques locales servant à calculer le pas $\delta\mathbf{p}$ soient en désaccord avec les variations effectives $\delta f_{\mathbf{p}}(\delta\mathbf{p})$. D'autant plus que lorsque la hessienne n'est pas positive définie on montre que δf^{quad} n'est pas une fonction bornée vers le bas, ce qui conduit à faire des pas de longueur infinie. C'est pourquoi on utilise

un algorithme itératif, qui varie une borne de longueur maximum de pas, en fonction de l'accord quadratique, et qui pour une borne donnée impose que la longueur du pas lui soit inférieure.

Le problème de région de confiance, pour une borne Δ donnée, est généralement exprimé sous la forme

$$\min_{\delta \mathbf{p}} \delta f_{\mathbf{p}}^{\text{quad}}(\delta \mathbf{p}) \quad \text{tq } \|\delta \mathbf{p}\|^2 \leq \Delta. \quad (2.47)$$

On démontre [Moré and Sorensen, 1983] que cette formulation est équivalente à

$$\min_{\delta \mathbf{p}} \delta f_{\mathbf{p}}^{\text{quad}}(\delta \mathbf{p}) + \ell \|\delta \mathbf{p}\|^2 \quad (2.48)$$

avec ℓ ajusté de telle sorte que

$$\begin{aligned} \ell &= 0 & \text{et } \|\delta \mathbf{p}\| &\leq \Delta \\ \text{ou } \ell &> 0 & \text{et } \|\delta \mathbf{p}\| &= \Delta \end{aligned} \quad (2.49)$$

En utilisant une norme quadratique $\|\delta \mathbf{p}\|_{\mathbf{D}}^2 = \delta \mathbf{p}^T \cdot \mathbf{D} \cdot \delta \mathbf{p}$ dans l'équation (2.48), avec \mathbf{D} positive définie, on obtient le pas de région de confiance (Trust-Region) :

$$\delta \mathbf{p}^{\text{TR}} = -(\mathbf{A} + \ell \mathbf{D})^{-1} \cdot \mathbf{g}. \quad (2.50)$$

La mise à jour de la borne Δ de la région de confiance se fait à chaque itération comme il suit. La borne Δ est réduite et le pas est recalculé en fonction du nouveau Δ , lorsque le rapport entre les variations de la fonction à minimiser et de son approximation quadratique $\rho < \zeta_1 < 1$ est trop petit ou $\rho > \zeta_4 > 1$ est trop grand, avec typiquement $\zeta_1 \approx 0.1$ et $\zeta_4 \approx 10$. Le pas est accepté et Δ reste inchangée, lorsque l'accord quadratique est correct, avec $\zeta_1 \leq \rho < \zeta_2 < 1$ ou $1 < \zeta_3 < \rho \leq \zeta_4$ (voir schéma 2.2). Enfin, le pas est accepté et

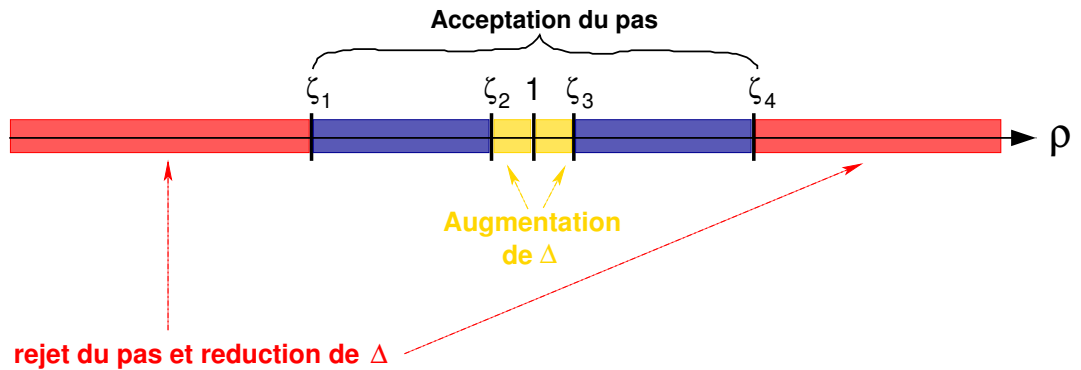


FIG. 2.2 – Mise à jour de la région de confiance

la taille de la borne est accrue, lorsque l'accord quadratique est excellent, c'est-à-dire ρ très proche de 1, avec $\zeta_2 \leq \rho \leq \zeta_3$. Typiquement $\zeta_2=0.95$ et $\zeta_3=1.05$.

La région de confiance est particulièrement adaptée pour contrôler les pas dans les algorithmes de quasi-Newton, car les pas sont dérivés directement et exclusivement d'après les approximations quadratiques, et car elle aide à pallier les problèmes de mauvais conditionnement. De plus sa mise en œuvre est également particulièrement simple et automatisée, et elle se prête à des développements analytiques moins complexes que les autres méthodes d'ajustement de longueur de pas. Toutefois elle requiert que la matrice hessienne soit suffisamment bien approximée, avec le bon facteur scalaire de remise à l'échelle notamment.

2.6 Exploration discontinue de type Monte-Carlo

2.6.1 Débuts historiques

De façon générale, les méthodes de Monte-Carlo désignent des techniques d'échantillonnages aléatoires visant à résoudre des problèmes mathématiques numériquement compliqués. L'exemple pittoresque faisant date de première expérience de Monte-Carlo est généralement attribué à Buffon [1777]. Celui-ci avait en effet mis en place une expérience aléatoire de jet d'aiguilles de longueur l sur un plancher en bois avec des fissures régulières à pas s , permettant d'obtenir une estimation de π à partir de la probabilité que m aiguilles aient croisé une fissure au bout de N réalisations. Il avait en effet préalablement calculé que $\frac{m(N)}{N} \xrightarrow{N \rightarrow \infty} \frac{2}{\pi} \frac{s}{l}$. Le développement de cette approche fut relancé à peu près cent ans après, par la découverte que les intégrales à grands nombres de paramètres \mathbf{p} , de type

$$I = \int_{\mathcal{P}} f(\mathbf{p}) \ell(\mathbf{p}) d\mathbf{p}, \quad (2.51)$$

où $\ell(\mathbf{p})$ est intégrable et à valeurs réelles positives, peuvent être estimées par

$$\tilde{I} = \frac{1}{N} \sum_{i=1}^N f(\mathbf{p}_i). \quad (2.52)$$

où les \mathbf{p}_i sont des réalisations de la densité de probabilité $\ell(\mathbf{p}) / \int_{\mathcal{P}} \ell(\mathbf{p}) d\mathbf{p}$. Mais il manquait alors de puissance de calcul numérique pour tirer profit de cette découverte pour le calcul intégrale.

2.6.2 Application aux problèmes inverses

Le premier emploi de méthodes Monte-Carlo dans un contexte de problème inverse fut effectué par le groupe de chercheurs associés à Métropolis, pour le projet Los

Alamos de réalisation de la première bombe nucléaire en 1940 [Metropolis and Ulam, 1949; Metropolis et al., 1953]. Il fallait alors estimer des paramètres macroscopiques à partir de systèmes thermodynamiques à très grands nombres de degrés de liberté. Les travaux de Hastings [1970] puis Geman and Geman [1984] et le développement du calcul numérique avec les ordinateurs ont depuis permis d'étendre ces méthodes plus largement. Aujourd'hui, les méthodes d'explorations aléatoires de type *Monte-Carlo* sont en effet employées pour résoudre les problèmes inverses non-linéaires à la place — voir en complément — des méthodes d'optimisation pseudo-continues, dans les cas où le problème inverse à résoudre est trop fortement non-linéaire ou trop multi-modal ; lorsque les incertitudes sur les adéquations de $f_{\text{post}}(\mathbf{p})$ sont très loins d'être gaussiennes ; et lorsque les paramètres du problème inverse sont des variables discrètes.

Les méthodes *Monte-Carlo* sont en effet les méthodes aléatoires les plus rapides pour échantillonner les densités de probabilité et en extraire des critères intégraux [Mosegaard and Sambridge, 2002], lorsque la dimension $M_{\mathcal{P}}$ de l'ensemble des paramètres \mathcal{P} est très élevée et que les zones de forte densité de probabilité correspondent en conséquence à un volume très raréfié des paramètres. Par exemple on voit que la stratégie aléatoire la plus simple d'échantillonnage de $f_{\text{post}}(\mathbf{p})$, qui consiste à tirer uniformément des paramètres \mathbf{p}^{prop} et à les accepter/rejeter en fonction de la valeur effective de $f_{\text{post}}(\mathbf{p}^{\text{prop}})$, devient très vite impraticable avec l'augmentation de $M_{\mathcal{P}}$.

Quelques articles font état d'une possible utilisation de techniques Monte-Carlo pour la reconstruction de phase [Nieto-Vesperinas and Mendez, 1986; Nieto-Vesperinas et al., 1988; Burian et al., 2000; Taylor et al., 2006]. Ce n'est toutefois pas l'approche que j'ai suivie dans les chapitres suivants, puisque les résultats obtenus par optimisation pseudo-continue sont actuellement plus convaincants dans le contexte de la reconstruction de phase, tout en nécessitant beaucoup moins de calcul. Néanmoins la combinaison de techniques Monte-Carlo à des techniques pseudo-continues, me semble constituer une possible issue pour augmenter encore le nombre de degrés de liberté par rapport aux résultats que j'ai obtenus dans les chapitres 3 et 4, et c'est pourquoi j'évoque quand même quelques principes de l'exploration Monte-Carlo pour les problèmes inverses.

2.6.3 Chaînes de Markov

La première grande famille d'algorithmes de type Monte-Carlo utilisée pour les problèmes inverses est constituée des méthodes de type MCMC (Monte-Carlo par chaînes de Markov). Elle sert principalement à extraire des critères intégraux de $f_{\text{post}}(\mathbf{p})$, tels que la moyenne a posteriori ou le maximum marginal a posteriori.

Les chaînes de Markov sont des séquences de variables aléatoires (processus stochastiques) vérifiant la propriété de Markov

$$f(\mathbf{p}^{(n+1)}|\mathbf{p}^{(n)}) = f(\mathbf{p}^{(n+1)}|\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \dots, \mathbf{p}^{(n)}), \quad (2.53)$$

selon laquelle la mémoire du cheminement précédent qui est utilisée pour déterminer l'exploration future, est limitée à la dernière exploration effectuée. Le choix de cette propriété sert principalement pour le problème inverse à limiter au maximum le stockage. Certains algorithmes utilisent toutefois la propriété à l'ordre k c'est-à-dire en gardant les k derniers pas en mémoires, ce qui se traduit aussi par une complexité accrue de l'algorithme.

2.6.3.1 Algorithme de Métropolis

L'algorithme de Métropolis est défini par rapport à l'algorithme d'acceptation/rejet pour échantillonner plus rapidement la distribution $f(\mathbf{p})$, en acceptant statistiquement comme nouvel échantillon à l'itération $(n+1)$, celui de $\mathbf{p}^{(n)}$ et d'un nouvel échantillon \mathbf{p}^{new} , qui a la densité de probabilité a posteriori f la plus significative. L'échantillon \mathbf{p}^{new} est proposé à l'aide d'une fonction aléatoire l , appelée loi instrumentale ou proposition, par

$$\mathbf{p}^{\text{new}} = l(\mathbf{p}^{(n)}). \quad (2.54)$$

Les conditions pour que l'algorithme fonctionne sont que la loi l doit échantillonner une densité de probabilité constante au bout d'un certain nombre d'itérations, et être irréductible et apériodique, c'est-à-dire que tous les ensembles de probabilité non nulle peuvent être atteints à partir de tout point de départ, et qu'aucun noyau n'induit de comportement périodique dans les échantillons proposés.

En notant $t^{(n+1)}$ un tirage aléatoire uniforme sur $[0, 1]$ permettant d'introduire dans la comparaison $f(\mathbf{p}^{\text{new}})$ et $f(\mathbf{p}^{(n)})$, une notion statistique qui est nécessaire à une certaine évolution de l'échantillonnage, le nouvel échantillon est ainsi déterminé par

$$\mathbf{p}^{(n+1)} = \begin{cases} \mathbf{p}^{\text{new}} & \text{si } t^{(n+1)} \leq \min \left[1, \frac{f(\mathbf{p}^{\text{new}})}{f(\mathbf{p}^{(n)})} \right] \\ \mathbf{p}^{(n)} & \text{sinon} \end{cases} \quad (2.55)$$

Une extrapolation de cet algorithme initial, connue sous le nom d'échantillonneur de Gibbs, a été proposée par Geman and Geman [1984]. Elle consiste à proposer plusieurs nouveaux échantillons à chaque itération au lieu d'un seul.

Concrètement pour échantillonner f_{post} , on peut choisir une fonction l , qui échantillonne f_{prior} lorsqu'elle est appliquée successivement, et $f = f_{\text{exp}}$. De plus, lorsque

$$f_{\text{exp}}(\mathbf{p}) = f_{\text{exp}_1}(\mathbf{p}) \times f_{\text{exp}_2}(\mathbf{p}) \dots \times f_{\text{exp}_n}(\mathbf{p}), \quad (2.56)$$

comme par exemple dans le cas de la diversité de phase, on peut utiliser un algorithme cascadié de Métropolis [Mosegaard, 2006], dans lequel une loi l échantillonnant $f_{\text{prior}} \times \prod_{k=0}^{i-1} f_{\text{exp}_k}$ est utilisé comme loi de proposition pour échantillonner $f_{\text{prior}} \times \prod_{k=0}^i f_{\text{exp}_k}$.

2.6.4 Recuit simulé

Le recuit-simulé est quand à lui utilisé dans le cadre des problèmes inverses, pour trouver les paramètres qui maximisent f_{post} — en pratique ceux qui minimisent son logarithme négatif f' . Il est inspiré de l'algorithme de Métropolis et des techniques de minimisation de l'énergie d'un matériau en métallurgie par la répétition d'un cycle alternant un lent refroidissement et un réchauffage rapide (pour éliminer les impuretés). Il a été proposé par Kirkpatrick et al. [1983] pour la première fois en tant que “simulated annealing”, c'est-à-dire avec un cycle de lent refroidissement uniquement. Il est basé sur la densité de probabilité de Gibbs-Boltzmann

$$g_T(\mathbf{p}) = \frac{\exp(-E(\mathbf{p})/T)}{\text{cste}(T)}, \quad (2.57)$$

qui détermine la distribution de particules d'un système thermodynamique en équilibre en fonction de son énergie.

Les variations de température du recuit-simulé permettent de prendre un chemin qui descend progressivement la température vers une température nulle, malgré les minimums locaux (les impuretés du matériau). A température nulle, $g(\mathbf{p})$ approxime en effet une distribution delta où l'énergie $E(\mathbf{p})$ a été amenée à son minimum global.

En pratique l'algorithme consiste à combiner une stratégie de variation de température, avec des itérations qui consistent chacune en la proposition d'un pas $\delta\mathbf{p}$ et son acceptation avec la probabilité $g_T(\delta\mathbf{p})$.

La recommandation initiale de Kirkpatrick est de diminuer la température exponentiellement, toutefois cette stratégie ne fonctionne pas sur beaucoup de cas difficiles [Pedersen, 1990], et mène alors vers un minimum local. Une autre stratégie de refroidissement consiste à diminuer la température à chaque fois qu'un équilibre est atteint. Toutefois cela mène à une optimisation avec des temps infinis. De nombreux travaux sur le recuit-simulé dans les années 80 dont ceux de Rothman sur des données sismiques ont révélé la difficulté du réglage de la température, et donc la difficulté d'utilisation pratique de l'algorithme. Cependant des travaux plus prometteurs ont par la suite consisté à diminuer la température progressivement en se basant sur une diminution de température à distance constante de l'équilibre [Nulton and Salamon, 1988; et al., 1988].

Il resterait encore de nombreuses techniques de Monte-Carlo à évoquer telles que notamment les codes génétiques et les réseaux neuronaux, mais je m'arrête ici dans cette présentation, puisque les développements que je présente par la suite ne mettent pas en œuvre de technique Monte-Carlo.

Chapitre 3

Reconstruction de phase à partir d'une image tavelée

Dans ce chapitre je présente mon travail sur l'estimation des aberrations atmosphériques de la phase sur la pupille d'un télescope, à partir des images tavelées au foyer du télescope. C'est un problème inverse non-linéaire difficile à grand nombre de paramètres dont je rappelle tout d'abord l'historique. Je présente ensuite mon approche pour le contexte particulier auquel je me suis intéressé de l'estimation des fortes aberrations (D/r_0 élevés). En particulier je présente un nouvel algorithme d'optimisation global pour résoudre ce problème pour des conditions de turbulence plus difficiles (D/r_0 plus grands) que ce qui a été fait jusqu'à présent. Les performances de l'algorithme sont présentées avec de nombreuses simulations en fonction du D/r_0 et du niveau de bruit.

3.1 Historique

La reconstruction de phase à partir de mesures d'intensités est un problème inverse récurrent en physique. La microscopie électronique, la cristallographie, l'optique active, la déconvolution aveugle, la reconstruction d'images 3D par tomographie, ont notamment motivé son développement au cours des 50 dernières années. Les premiers algorithmes permettant de mesurer des petites aberrations à partir d'une image dans un plan focal ont été proposés par Gerchberg and Saxton [1972], puis Gonsalves [1976] et Fienup [1978]. Ces premiers algorithmes sont basés sur la variation itérative de l'amplitude complexe dans les deux plans de Fourier associés de la pupille et de la mesure. A chaque itération l'amplitude complexe est mise à jour dans son plan grâce respectivement à la contrainte de support de la pupille et grâce aux contraintes associées aux mesures. Fienup [1982] a ensuite établi la comparaison entre ces algorithmes et les algorithmes d'optimisation locale de type plus grande pente et gradients conjugués. Il

a également montré que ces premières méthodes possèdent les mêmes propriétés de convergence locale au voisinage d'un point fixe, que les algorithmes d'optimisation pseudo-continus de la comparaison.

Pour pallier la stricte ambiguïté de signe dans la reconstruction de phase, Gonsalves [1982] a proposé la technique de diversité de phase, qui s'est rapidement imposée. Cette technique permet de démanteler l'ambiguïté grâce à une deuxième image, contenant par rapport à la première une aberration fixe d'ordre pair, qui est généralement une défocalisation. L'ambiguïté peut toutefois être levée sans diversité dans le cas d'acquisitions continues, grâce à la corrélation temporelle entre les phases successives, ce que je montre à la section 3.8. Un état de l'art des méthodes de reconstruction de phase avec diversité focale a récemment été écrit par Mugnier et al. [2006].

Dans le contexte particulier auquel je me suis intéressé d'estimation des aberrations atmosphériques de phase à partir d'une seule image tavelée au foyer d'un télescope, les développements de référence sont ceux accomplis par Irwan and Lane [1998]. Ils ont en effet obtenu la convergence locale de la phase pour des rapports D/r_0 jusqu'à 4 en faisant l'optimisation par gradients conjugués du critère de maximum de vraisemblance de l'adéquation a posteriori (critère MAP). Je démontre par la suite qu'on arrive à reconstruire la phase pour des conditions de turbulence plus difficiles que $D/r_0=4$, en partant d'une formulation analogue et en utilisant un algorithme d'optimisation non-linéaire mieux adapté, que je combine à une stratégie d'optimisation globale efficace lorsque l'optimisation locale ne suffit plus. A ma connaissance, mon algorithme est le premier algorithme d'optimisation globale effectif dans ce contexte de reconstruction de phase sans utiliser de diversité de phase.

3.2 Adéquation expérimentale (ML)

3.2.1 Modèle direct des images

Le modèle continu reliant les images au foyer d'un télescope à l'amplitude complexe sur la pupille du télescope, a été explicité à la section 1.1.2.2. La discrétisation se fait de la manière suivante :

$$\begin{aligned} p_k &= P(\lambda \mathbf{u}_k) \\ \varphi_k &= \Phi(\lambda \mathbf{u}_k) \\ m_j &= \mathcal{M}(\mathbf{x}_j) \end{aligned} \tag{3.1}$$

où p_k est la pupille échantillonnée, φ_k est la phase échantillonnée et m_j est le modèle des observations au j -ième pixel. \mathbf{x} représente la position spatiale en unité angulaire sur le détecteur placé dans le plan focal, et \mathbf{u} la fréquence spatiale qui lui est associée. Le modèle discret de l'image s'écrit

$$\mathbf{m}(\alpha, \boldsymbol{\varphi}) = \alpha |\mathbf{F} \cdot \mathbf{a}(\boldsymbol{\varphi})|^2, \tag{3.2}$$

où $\alpha > 0$ rend compte du flux dans l'image, \mathbf{F} est l'opérateur linéaire de Fourier discret, et $\mathbf{a}_\varphi = \mathbf{p} \times \exp(i\varphi)$ est l'amplitude complexe de l'onde EM sur la pupille. Dans cette dernière expression le symbole \times représente la multiplication point à point, et la notation \exp est utilisée pour désigner le nouveau vecteur obtenu par l'application de l'exponentielle à chacune des composantes du vecteur d'entrée. L'omission de l'indice de longueur d'onde dans l'équation (3.2) est faite pour simplifier les écritures, mais cette relation est néanmoins fortement monochromatique.

L'hétérogénéité de sensibilité à l'intérieur des pixels du détecteur n'est pas prise en compte dans mon modèle discret. En général quand on prend en compte ces effets, on suppose que la sensibilité d'un pixel à l'autre est à peu près la même, et on convolue les échantillons de l'image par une réponse spatiale moyenne des pixels. Par ailleurs, le modèle que j'ai décliné décrit l'observation d'une source ponctuelle. Il doit donc également faire l'objet d'une convolution par la distribution angulaire d'énergie de la source, lorsque la source observée est résolue par le télescope. L'algorithme d'optimisation globale qui suit peut cependant être utilisé à l'identique lorsque les distributions qui convoluent le modèle sont connues. Il nécessiterait quelques légères retouches si les distributions doivent être conjointement estimées et si elles sont ajoutées dans le problème inverse sous la forme d'une déconvolution aveugle a posteriori.

Les paramètres de la reconstruction de phase sont $\mathbf{p} = \{\alpha, \varphi\}$. Dans la suite j'ai utilisé diverses reparamétrisations linéaires de la phase pour favoriser la convergence des algorithmes d'optimisation, et j'ai séparé les modes de tip-tilt de la phase. Celle-ci se réécrit donc

$$\varphi = \mathbf{B} \cdot \boldsymbol{\theta} + \mathbf{K} \cdot \boldsymbol{\phi}, \quad (3.3)$$

où \mathbf{K} est l'opérateur linéaire de reparamétrisation de la phase, $\mathbf{B} = \frac{2\pi}{N} [\mathbf{u}_1, \mathbf{u}_2]$ est la base de tip-tilt en unité de déplacement dans l'image, N est le nombre de pixels et $\mathbf{u}_1, \mathbf{u}_2$ sont les positions selon chaque direction du plan décrit par \mathbf{u} . Les paramètres de la reconstruction de phase se réexpriment donc en fonction des coefficients de tip-tilt $\boldsymbol{\theta}$ et des coefficients $\boldsymbol{\phi}$ de la nouvelle base :

$$\mathbf{p} = \{\alpha, \boldsymbol{\theta}, \boldsymbol{\phi}\}. \quad (3.4)$$

3.2.2 Bruits et incertitudes expérimentales

Le détecteur (CCD) reçoit une énergie quantifiée fluctuante. On ne peut donc pas compter à flux moyen constant la même énergie pour deux intervalles de temps égaux et séparés. On montre en physique quantique que la probabilité de détection de K photons pour un nombre moyen de N photons suit la loi de Poisson. Cette fluctuation est inévitable, dépend du flux moyen et doit être prise en compte pour chaque pixel du modèle. La variance associée à ce bruit s'écrit $\sigma_{\text{ph}_j}^2 = N_j$, où N_j est le nombre moyen de photons pour le pixel j de l'image.

Le détecteur introduit des fluctuations supplémentaires sur le signal photonique. Je néglige dans le modèle, comme on l'a vu, l'hétérogénéité de transmission des pixels, et je considère que le détecteur est uniquement responsable d'un bruit additif gaussien centré et spatialement homogène de variance σ_{CCD}^2 . Celui-ci correspond à la somme des bruits de lecture et de digitalisation principalement, et est généralement considéré homogène spatialement pour un CCD.

L'incertitude expérimentale due à la modélisation est négligée dans mes simulations, mais peut être prise facilement en compte le cas échéant sans modifier les méthodes développées.

Lane [1996] a montré que l'approximation globale des fluctuations gaussiennes sur l'adéquation expérimentale est une approximation robuste pour obtenir le maximum de vraisemblance, lorsque les données sont affectées par le bruit de Poisson et un bruit gaussien. Cette approximation qui est intuitive à fort flux, lorsque la distribution de Poisson tend vers celle de Gauss, est également valable pour des flux plus modérés. C'est donc l'hypothèse que j'ai retenue en écrivant l'adéquation expérimentale sous la forme $\mathbf{d} = \mathbf{m} + \mathbf{b}$ avec $\sigma_{\mathbf{b}_j}^2 = N_{\text{ph}_j} + \sigma_{\text{CCD}}^2$. En pratique le nombre de photons N_{ph_j} est déduit de l'image. Il est pris égal à 0 lorsque l'intensité de l'image sur le pixel j est négative à cause de la soustraction du champ obscur du CCD.

3.2.3 Adéquation expérimentale

Sous l'hypothèse des fluctuations gaussiennes entre le modèle et l'image acquise, la fonction d'adéquation expérimentale à minimiser est prise à une constante multiplicative près, comme le logarithme négatif de la densité de probabilité expérimentale, et s'écrit

$$f_{\text{exp}}(\mathbf{p}) = \frac{1}{2} (\mathbf{d} - \mathbf{m}(\mathbf{p}))^\top \cdot \mathbf{C}_{\mathbf{b}}^{-1} \cdot (\mathbf{d} - \mathbf{m}(\mathbf{p})) , \quad (3.5)$$

où $\mathbf{C}_{\mathbf{b}} = \langle \mathbf{b} \cdot \mathbf{b}^\top \rangle$ est la matrice de covariance spatiale des erreurs. De plus dans le cas simplifié où les bruits considérés sont spatialement décorrélés, la matrice de covariance est diagonale avec $C_{\mathbf{b}_{jj}} = \sigma_{\mathbf{b}_j}^2$ et f_{exp} s'écrit comme une fonction de *moindres carrés* :

$$f_{\text{exp}}(\mathbf{p}) = \frac{1}{2} \sum_j \left(\frac{d_j - m_j(\mathbf{p})}{\sigma_j} \right)^2 . \quad (3.6)$$

3.3 Adéquations a priori

3.3.1 Ajustement analytique non-ambigu du paramètre d'amplitude

Le paramètre d'amplitude α est ajusté de façon indépendante des paramètres de phase $\{\theta, \phi\}$, par une expression analytique ne contenant aucune ambiguïté. Il ne né-

cessite donc pas de régularisation, contrairement aux autres paramètres. Je lui règle donc ici son cas avant de discuter des ambiguïtés des autres paramètres et des a priori choisis pour remédier à leurs ambiguïtés.

Le modèle des images en fonction de l'amplitude α s'écrit

$$\mathbf{m}(\alpha, \boldsymbol{\phi}) = \alpha \mathbf{q}(\boldsymbol{\phi}), \quad (3.7)$$

avec $\mathbf{q}(\boldsymbol{\phi}) = |\mathbf{F} \cdot \mathbf{a}(\boldsymbol{\phi})|^2$. La fonction expérimentale f_{exp} est également linéaire par rapport au paramètre α et peut ainsi se réécrire indépendamment de α pour l'optimisation, sous la forme simplifiée

$$f_{\text{exp}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{1}{2} \sum_j \left(\frac{d_j - \alpha^+ q_j(\boldsymbol{\theta}, \boldsymbol{\phi})}{\sigma_j} \right)^2, \quad (3.8)$$

où le paramètre α^+ est ajusté indépendamment de la phase à chaque itération, par optimisation linéaire avec $(\boldsymbol{\theta}, \boldsymbol{\phi})$ fixés :

$$\alpha^+ = \min_{\alpha} \left[\sum_j \left(\frac{d_j - \alpha q_j}{\sigma_j} \right)^2 \right] = \frac{\sum_j (d_j \times q_j) / \sigma_j^2}{\sum_j (q_j \times q_j) / \sigma_j^2} \quad (3.9)$$

Le nouveau jeu de paramètres considéré dans toute la suite est donc $\mathbf{p} = \{\boldsymbol{\theta}, \boldsymbol{\phi}\}$ et le modèle n'est plus exprimé en fonction de l'amplitude mais simplement par $\mathbf{m}_{\mathbf{p}} = \mathbf{q}_{\mathbf{p}}$. De plus on montre par chaînage sur les dérivées partielles, que les dérivées de la nouvelle fonction $f_{\text{exp}}(\mathbf{p}) \triangleq f_{\text{exp}}(\alpha^+(\mathbf{p}), \mathbf{p})$, utilisées pour obtenir les pas de l'optimisation itérative des paramètres \mathbf{p} , ne dépendent plus de α . En effet en utilisant que $\frac{\partial f_{\text{exp}}(\mathbf{p})}{\partial \alpha} \big|_{\alpha=\alpha^+(\mathbf{p})} = 0$ d'après l'équation (3.9), on obtient

$$\begin{aligned} \frac{\partial f_{\text{exp}}(\mathbf{p})}{\partial p_k} &= \frac{\partial f_{\text{exp}}(\alpha, \mathbf{p})}{\partial p_k} \big|_{\alpha=\alpha^+(\mathbf{p})} + \frac{\partial \alpha^+(\mathbf{p})}{\partial p_k} \times \frac{\partial f_{\text{exp}}(\mathbf{p})}{\partial \alpha} \big|_{\alpha=\alpha^+(\mathbf{p})} \\ &= \frac{\partial f_{\text{exp}}(\alpha, \mathbf{p})}{\partial p_k} \big|_{\alpha=\alpha^+(\mathbf{p})}. \end{aligned} \quad (3.10)$$

3.3.2 Ambiguïtés de la phase dans l'adéquation expérimentale

3.3.2.1 Ambiguïté stricte du signe de la composante paire de la phase

L'image au foyer du télescope est proportionnelle à la densité spectrale de l'amplitude complexe $a_{\varphi}(\mathbf{u})$ sur la pupille. Ainsi d'après le théorème de Wiener-Khinchine, elle peut être alternativement considérée comme proportionnelle à la transformée de Fourier de l'autocorrélation :

$$\Gamma_{a_{\varphi}}(\boldsymbol{\tau}) \triangleq \iint_{\mathcal{P}} a_{\varphi}(\mathbf{u} + \boldsymbol{\tau}) a_{\varphi}^*(\mathbf{u}) d\mathbf{u}. \quad (3.11)$$

En notant $a'_\varphi(\mathbf{u}) = a_{-\varphi}^*(-\mathbf{u})$, on montre [Blanc, 2002] que $\Gamma_{a'_\varphi} = \Gamma_{a_\varphi}$. Si la pupille est symétrique et réelle alors $a'_\varphi(\mathbf{u}) = P(\mathbf{u}) e^{-i\varphi(-\mathbf{u})}$, et l'on obtient que les deux phases $\varphi(\mathbf{u})$ et $-\varphi(-\mathbf{u})$ correspondent strictement aux mêmes images. De plus en décomposant la phase en parties paire et impaire :

$$\begin{aligned}\varphi(\mathbf{u}) &= \frac{\varphi(\mathbf{u}) + \varphi(-\mathbf{u})}{2} + \frac{\varphi(\mathbf{u}) - \varphi(-\mathbf{u})}{2} \\ &= \varphi_p(\mathbf{u}) + \varphi_i(\mathbf{u}),\end{aligned}\tag{3.12}$$

on voit que $\varphi(\mathbf{u})$ et $-\varphi(-\mathbf{u})$ ont la même partie impaire et une partie paire opposée. L'ambiguïté entre les deux phases $\varphi(\mathbf{u})$ et $-\varphi(-\mathbf{u})$ dans les images correspond donc à une ambiguïté du signe de la partie paire de φ .

3.3.2.2 Ambiguïté de signe de chaque mode pair de la phase

Plus généralement si on décompose la phase sur une base modale et qu'on considère que les modes reconstruits ont à peu près la bonne statistique grâce à la régularisation, leurs effets dans l'image sont plus difficilement portés par d'autres modes. On voit alors que le signe d'un mode pair a des chances de mener à un minimum local si son poids est important dans la phase. En pratique j'ai en effet pu constater en affichant ces modes pairs lors de la reconstruction de la phase, que certaines phases reconstruites sont identiques à la phase recherchée, au signe près de quelques premiers modes. Dans ce cas l'adéquation expérimentale est toutefois moins bonne que si la phase a les bons signes, car les ambiguïtés ne sont pas strictes.

3.3.2.3 Ambiguïtés strictes de périodicité modulo 2π de la phase

L'exponentielle complexe dans le modèle des images, implique que l'ajout d'une période à un échantillon de phase, ne change pas le modèle de l'image. Cet ajout ne change donc pas non plus l'adéquation expérimentale f_{exp} servant à estimer les paramètres de phase :

$$f_{\text{exp}}(\mathbf{p}_{i \neq j}, \varphi_j + k 2\pi) = f_{\text{exp}}(\mathbf{p}_{i \neq j}, \varphi_j).\tag{3.13}$$

Ce sont ces ambiguïtés qui sont les plus gênantes dans le cadre de la reconstruction de phases à fortes amplitudes. Si on souhaite appliquer une correction de phase a posteriori de type déconvolution aveugle de l'image, on peut les ignorer et rechercher uniquement la phase pliée, c'est à dire la phase dans l'intervalle $]-\pi, \pi]$. Cette approche a été récemment menée par Mugnier et al. [2006]. Toutefois ce n'est pas l'approche adéquate lorsqu'on veut corriger les aberrations de phase pour n'importe quelle longueur d'onde dans le plan pupille avec une optique adaptative. Dans le cadre ELP-OA de reconstruction de phase avec diversité polychromatique, notamment, la paramétrisation commune en différence de marche implique d'estimer la phase dépliée pour avoir une

bonne estimation de la phase à toutes les longueurs d'onde. J'ai donc opté pour la recherche de la phase dépliée, y compris dans le cas monochromatique, en comptant sur les a priori pour déplier la phase, puisque le critère expérimental ne le permet pas du tout.

On remarque également que la valeur moyenne de la phase sur la pupille (pour une pupille à transmission uniforme) ou piston ne change pas le modèle des images. Pour lever cette ambiguïté de constante globale de la phase qui ne peut pas être estimée, j'ai choisi de reconstruire la phase à piston nul.

3.3.3 Adéquation a priori des paramètres de phase dépistonnée et détiltée

3.3.3.1 Densité de probabilité a priori

Pour démanteler les ambiguïtés de la phase et faciliter l'optimisation très non-linéaire de l'adéquation expérimentale, on se sert des a priori sur la statistique spatiale de la phase. En faisant l'hypothèse d'une turbulence de statistique Kolmogorov, Papoulis [1965] a montré que la phase dépistonnée suit une statistique gaussienne. Les modes de tip-tilt suivant également une statistique gaussienne [Tatarskii, 1971], il en est donc finalement de même de la phase dépistonnée et détiltée ; si bien que sa régularisation, en considérant le logarithme négatif s'écrit

$$f_{\text{prior}}(\phi) = \frac{1}{2} \phi^\top \cdot \mathbf{C}_\phi^{-1} \cdot \phi, \quad (3.14)$$

aux constantes additives et multiplicatives inutiles près, et où $\mathbf{C}_\phi = \langle \phi \cdot \phi^\top \rangle$ est la matrice de covariance de la phase dépistonnée et détiltée. Le détiltage et le dépistonnage de ϕ peuvent être modélisés et codés comme l'application d'un projecteur symétrique \mathbf{P} :

$$\phi = \mathbf{P} \cdot (\mathbf{P} \cdot \varphi) = \mathbf{P} \cdot \varphi, \quad (3.15)$$

avec $\mathbf{P} = \mathbf{I} - (\mathbf{u}_0 \cdot \mathbf{u}_0^\top / \|\mathbf{u}_0\|^2 + \mathbf{u}_1 \cdot \mathbf{u}_1^\top / \|\mathbf{u}_1\|^2 + \mathbf{u}_2 \cdot \mathbf{u}_2^\top / \|\mathbf{u}_2\|^2)$, où \mathbf{u}_0 est le vecteur composé de 1 sur le support de la pupille, et où \mathbf{u}_1 et \mathbf{u}_2 sont les vecteurs de basculement (tip-tilt) définis à l'équation (3.3). Ainsi la matrice de covariance dépistonnée et détiltée s'écrit simplement

$$\mathbf{C}_\phi = \mathbf{P} \cdot \mathbf{C}_\varphi \cdot \mathbf{P}^\top = \mathbf{P} \cdot \mathbf{C}_\varphi \cdot \mathbf{P}. \quad (3.16)$$

Dans le cas de la reparamétrisation par l'opérateur linéaire \mathbf{K} et en notant $\mathbf{T} = \mathbf{K}^{-1} \cdot \mathbf{P}$, elle s'écrit

$$\mathbf{C}_\phi = \mathbf{T} \cdot \mathbf{C}_\varphi \cdot \mathbf{T}^\top. \quad (3.17)$$

Cependant comme \mathbf{P} est de rang $N-3$, alors la matrice inverse de \mathbf{C}_ϕ , qui doit être utilisée pour la régularisation des paramètres ϕ à l'équation (3.14) est indéfinie. On peut alors définir sa matrice pseudo-inverse à partir de la matrice pseudo-inverse de

\mathbf{P} , en éliminant les directions inutiles au sens des moindres carrés. On peut également choisir de régulariser une phase φ' avec un piston et des modes de tip-tilt indifférents et éliminés par application du projecteur \mathbf{P} pour calculer le critère f_{exp} et ses dérivées. Cela implique cette fois-ci une matrice inverse qui est définie :

$$f_{\text{prior}}(\varphi') = \frac{1}{2} \varphi'^{\top} \cdot \mathbf{C}_{\varphi'}^{-1} \cdot \varphi' . \quad (3.18)$$

J'ai vérifié en pratique que la phase ϕ est tout aussi bien régularisée par les deux approches, et je présente ici la deuxième, par simplicité.

Sous l'hypothèse de variance uniforme, la matrice de covariance $\mathbf{C}_{\varphi'}$ de la phase pistonnée et tiltée est simplement reliée à la fonction structure par

$$\mathbf{C}_{\varphi'} = \text{cste} - \frac{1}{2} \mathbf{D}_{\varphi'} , \quad (3.19)$$

avec $(\mathbf{D}_{\varphi'})_{i,j} = 6.88 \left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{r_0} \right)^{5/3}$. Comme l'opérateur linéaire de régularisation $\mathbf{C}_{\varphi'}^{-1}$ ne change pas avec la variation des paramètres lors de l'optimisation, il peut être pré-calculé. De plus il présente également l'avantage de ne pas dépendre des conditions particulières de turbulence. En effet la variation de D/r_0 se traduit pour l'opérateur de régularisation par un facteur scalaire, qu'on peut faire porter par l'hyperparamètre μ du poids de la régularisation. Cette simplification est d'autant plus justifiée que le poids de la régularisation est en fait ajusté automatiquement dans l'algorithme, en fonction de la rugosité locale de l'adéquation expérimentale, plutôt que d'après la valeur théorique attendue lorsque la force relative de la turbulence est connue.

Lorsque le nombre de paramètres devient important, l'opérateur de régularisation $\mathbf{C}_{\varphi'}^{-1}$ doit être appliqué rapidement pour ne pas paralyser les calculs de l'optimisation. Dans ce cas-là j'utilise une approximation creuse quasi-exacte de $\mathbf{C}_{\varphi'}^{-1}$, décrite dans [Béchet et al., 2006].

3.3.3.2 A priori alternatif de lissage

Comme les ambiguïtés de périodicité modulo 2π sont les ambiguïtés les plus importantes pour la reconstruction dans le cas de phases à fortes amplitudes, on peut penser à utiliser un simple a priori de lissage de la phase dans l'adéquation a posteriori. On a en effet intérêt à imposer pour chaque échantillon de phase, que les échantillons voisins aient la même périodicité, pourvu que le nombre de paramètres N_p soit suffisamment grand devant $(D/r_0)^2$ — i.e. pourvu que la phase soit échantillonnée suffisamment finement par rapport à la taille caractéristique de sous-pupille qui correspond à des fluctuations de 1 radian—, ce qui peut-être réalisé par un a priori du type

$$f_{\text{liss}}(\varphi) = \varphi^{\top} \cdot \mathbf{L} \cdot \varphi = \varphi^{\top} \cdot [(\mathbf{I} - \mathbf{S}) \cdot (\mathbf{I} - \mathbf{S})] \cdot \varphi , \quad (3.20)$$

où \mathbf{I} est la matrice identité et \mathbf{S} est un opérateur de lissage symétrique rapide, par exemple une convolution de l'image dans chaque direction par le noyau de convolution $[1/4, 1/2, 1/4]$ normalisé.

3.3.4 Ambiguïtés de tip-tilt

Le simple déplacement de la structure tavelée reconstruite se traduit par un recouvrement très multi-modal entre les tavelures (speckles) des images reconstruites et celles des données (voir la figure 5.5) et correspond donc à des fortes et nombreuses fluctuations du critère f_{exp} , avec de nombreux minima locaux. La valeur du tip-tilt est donc très ambiguë dans l'ajustement et déterminante pour assurer la bonne reconstruction de la phase. Comme on le voit à partir des résultats de la section 5.1.5.2 pour les forts D/r_0 , ce n'est pas une bonne idée d'ajuster le tip-tilt indépendamment de la phase par maximum de corrélation tant que la phase n'est pas suffisamment reconstruite. Il faut en effet suffisamment bien reconstruire la phase détiltée pour estimer le tip-tilt précisément et vice versa.

3.3.5 Régularisation du tip-tilt

Les paramètres de tip-tilt étant déterminants dans le succès de l'optimisation et très fluctuants d'une réalisation de phase à l'autre, j'utilise donc un bornage du tip-tilt d'après leur estimation dans les images. Le bornage peut s'exprimer sous la forme d'une régularisation :

$$R(\boldsymbol{\theta}) = \frac{1}{2} \frac{(\theta_1 - \widetilde{\theta}_1)^2 + (\theta_2 - \widetilde{\theta}_2)^2}{\sigma_{\text{tilt}}^2} \quad (3.21)$$

où $(\widetilde{\theta}_1, \widetilde{\theta}_2)$ est un estimateur du tip-tilt mesuré dans l'image, dont on suppose que l'erreur suit une statistique gaussienne et où σ_{tilt}^2 est la variance de cette erreur. Toutefois cela nécessite d'ajuster le poids relatif de $R(\boldsymbol{\theta})$ par rapport aux autres critères f_{exp} et f_{prior} . J'ai donc dans un premier temps imposé un bornage strict des coefficients de tip-tilt entre $\pm 3 \sigma_{\text{tilt}}$.

La variance σ_{tilt}^2 est en fait relativement importante. Le centroïde et le tip-tilt expriment en effet un déplacement assez différent des images tavelées (voir la section 5.1.1). Cette différence s'accroît fortement avec la force de la turbulence, pour devenir supérieure en moyenne à la distance entre deux speckles et donc ne plus permettre dans certains cas de contraindre la reconstruction de la phase dans le bon minimum local relativement aux variations exclusives de tip-tilt. Ce problème est en pratique responsable de l'actuelle limitation de l'algorithme à des $D/r_0 \leq 13$. Il nécessiterait de nouveaux développements, que je n'ai pas eu le temps d'accomplir, en optimisant par exemple un ensemble de phases à tilt fixé, et en regardant quelle valeur de tip-tilt permet d'obtenir la meilleure adéquation expérimentale.

3.4 Reparamétrisation

3.4.1 Importance de la reparamétrisation pour le problème inverse

Le choix des paramètres est souvent imposé lors de la formulation du problème inverse : dans notre cas on souhaite retrouver les paramètres de phase φ de l'amplitude complexe a_φ de l'onde EM sur la pupille du télescope. Cependant une reparamétrisation linéaire des paramètres recherchés à l'aide d'un opérateur \mathbf{K} , est généralement déterminante pour le succès de l'optimisation. Tout d'abord parce que le conditionnement de l'approximation \mathbf{A} de la hessienne $\nabla_{\mathbf{p}}^2 f_{\text{post}}$, utilisée pour calculer les pas $\delta \mathbf{p}$ successifs, peut significativement dépendre de la reparamétrisation (et donc également la convergence de l'algorithme d'optimisation). D'autre part parce que la reparamétrisation influe sur les ambiguïtés non strictes vues par l'ajustement, et qu'elle peut ainsi permettre de ne pas attribuer certaines variations du modèle aux mauvais paramètres.

Comme le modèle des images est non-linéaire par rapport à la phase, il n'existe pas de base linéairement reliée à la phase dont chaque mode puisse entraîner des effets orthogonaux dans les images par rapport à ceux des autres modes. Je ne connais par ailleurs pas de base, s'approchant plus que les autres de ce comportement idéal de base propre. J'ai toutefois comparé quelques différentes reparamétrisations de la phase (polynômes de Gram-Schmidt, de Zernike, de Karhunen-Loève, paramétrisations zonales avec différents échantillonneurs) pour voir si à défaut de pouvoir le démontrer, une de celle-ci est plus adaptée pour la reconstruction de phase. Mais ces tests ne m'ont pas non plus permis de choisir une reparamétrisation en particulier.

A défaut de l'orthogonalité, j'ai donc simplement opté pour reparamétriser la phase turbulente avec un nombre minimum de paramètres, ce qui m'a conduit à sélectionner deux bases en particulier, détaillées ci-dessous.

Notons enfin que l'amplitude α et les modes de tip-tilt θ ne font pas l'objet de la reparamétrisation, puisqu'ils sont considérés spécifiquement dans l'ajustement.

3.4.2 Reparamétrisation modale de la phase dépistonnée et détiltée

Il est d'usage commun en astronomie HRA, de reparamétriser la phase en utilisant les polynômes de Zernike, dont les tout premiers modes correspondent bien aux aberrations que l'on peut reconnaître à l'oeil dans les images de montages optiques (tip-tilt i.e. pente, défocalisation, coma).

Toutefois, comme démontré par Lane and Tallon [1992], la décomposition de la phase sur les polynômes de Zernike requiert un nombre de paramètres significativement plus important pour rendre compte de toutes les aberrations du front d'onde et donc modéliser les structures fines de l'image, que la paramétrisation en décomposant la phase sur les modes singuliers de la turbulence ou polynômes de Karhunen-Loève.

Notons également, comme on peut le voir sur la figure 3.1, que les premiers modes de ces derniers polynômes sont très semblables aux premiers modes de Zernike. Ils peuvent donc tout aussi bien modéliser les basses fréquences spatiales des images.

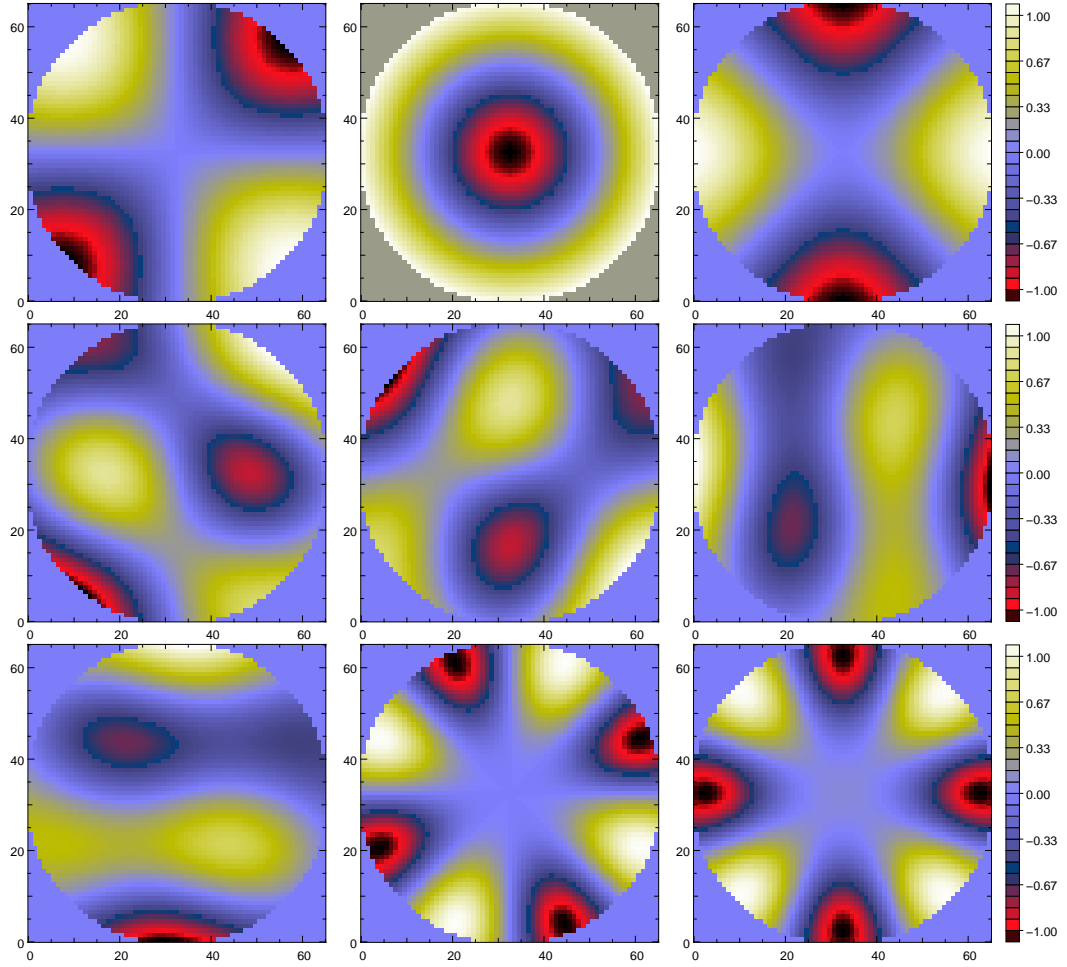


FIG. 3.1 – Représentation des premiers polynômes de Karhunen-Loève pour une statistique de la phase de type Kolmogorov

La matrice \mathbf{K}_{mod} de reparamétrisation modale de Karhunen-Loève de la phase dé-pistonnée et détiltée, est simplement obtenue par la décomposition numérique en valeurs singulières (SVD) de la matrice de covariance de Kolmogorov :

$$\mathbf{C}_\phi = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^\top \quad (3.22)$$

où \mathbf{U} est la matrice unitaire des vecteurs singuliers, \mathbf{V} est la matrice unitaire des vecteurs que l'endomorphisme correspondant à \mathbf{C}_ϕ transforme en vecteurs singuliers, et $\mathbf{\Sigma}$ est la matrice diagonale des valeurs singulières $\sigma_i > 0$. En effet on a directement

$\mathbf{K}_{\text{mod}} = \mathbf{U}$. De plus on remarque que cette décomposition correspond ici à une décomposition spectrale, avec $\mathbf{V}^T = \mathbf{U}^{-1}$, puisque \mathbf{C}_ϕ est symétrique.

3.4.3 Reparamétrisation zonale de la phase dépistonnée et détiltée

En considérant le rééchantillonnage des sous-paramètres de phase comme une opération de lissage, il vient assez naturellement l'idée d'utiliser un nombre restreint d'échantillons spatiaux de phase pour reparamétriser la phase alternativement. Cela permet en plus de la restriction du nombre de paramètres, de renforcer explicitement l'a priori de lissage de la phase et ainsi de faciliter le traitement des ambiguïtés strictes de périodicité de la phase dans l'adéquation f_{exp} .

J'ai choisi en particulier d'utiliser la reparamétrisation par les noeuds de fonctions splines linéaires bi-cubiques, et tous les résultats donnés par la suite ont été obtenus à l'aide de cette reparamétrisation, pour les premières itérations de l'algorithme assurant la convergence dans le bon minimum local tout au moins. La spline compacte d'ordre 3 utilisée a pour formule analytique :

$$\begin{aligned} \mathcal{S}(x) &= 2/3 + (x/2 - 1) \times x^2 & : & |x| \leq 1, \\ \mathcal{S}(x) &= 1/6 \times (2 - x)^3 & : & 1 \leq |x| \leq 2, \\ \mathcal{S}(x) &= 0 & : & |x| \geq 2. \end{aligned}$$

La matrice \mathbf{K}_{zon} de reparamétrisation contient la valeur aux échantillons de phase, des fonctions splines définies sur chaque noeud. \mathbf{K}_{zon} est donc un opérateur linéaire creux, qui est précalculé, stocké avec peu de mémoire et appliqué rapidement lors de son utilisation. Notons que cette paramétrisation à l'aide de splines est bien adaptée puisqu'on s'intéresse exclusivement à des cas où la grille de noeuds est moins échantillonnée que la grille de phase. Dans le cas contraire la description à l'aide de splines bi-cubiques n'est pas adaptée [Unser et al., 1995].

3.4.4 Restriction du nombre de paramètres

Il faut modéliser la phase de l'amplitude complexe sur la pupille avec de l'ordre de $(D/r_0)^2$ coefficients de la base de Karhunen-Loève ou bien des noeuds de splines, pour pouvoir reconstruire l'image jusqu'aux plus hautes fréquences spatiales de la turbulence. J'ai essayé de restreindre davantage ce nombre dans les premiers temps de l'optimisation, mais les paramètres ne s'ajustent alors pas bien, ce qui peut être en partie compris en se rappelant que les modes des bases utilisées n'ont pas des effets orthogonaux dans les images. Une bonne piste pour éviter ces ambiguïtés, mais que je n'ai pas eu le temps de mettre en pratique, me semble de binner l'image parallèlement à la restriction du nombre d'échantillons de la phase. En effet cela permettrait au moins que les hautes fréquences spatiales de l'image ne soient pas reconstruites par la phase sous-échantillonnée. Cette approche a d'ailleurs déjà été proposée dans un autre

contexte de reconstruction de phase [Won et al., 1985]. Au final la première étape de reconstruction de la phase est faite avec un nombre de paramètres restreints à de l'ordre de $(D/r_0)^2$.

Le nombre total d'échantillons de phase que j'utilise pour le calcul de l'image reconstruite et pour celui de l'image réelle simulée qui constitue les données \mathbf{d} , est au moins quinze fois supérieur (en 2D) à $(D/r_0)^2$. Cela afin d'une part de ne pas avoir de repliement dans les images, et d'autre part afin de renforcer le lissage de phase apporté par le rééchantillonneur et ainsi mieux gérer les ambiguïtés de périodicité de la phase (voir section 3.9.3).

3.5 Algorithme d'optimisation locale

3.5.1 Choix du type d'algorithme d'optimisation non linéaire

J'ai essayé divers algorithmes d'optimisation locale pour la reconstruction de phase, jusqu'à finalement développer un algorithme dédié. J'ai tout d'abord utilisé un algorithme de type quasi-Newton à métrique variable, puisque celui-ci est généralement très performant pour l'optimisation d'un grand nombre de paramètres, et ne nécessite à chaque itération que le calcul du gradient et le stockage de quelques vecteurs. De plus, Irwan and Lane [1998] ont obtenu des résultats probants dans le même contexte avec l'algorithme assez proche (et également testé) des gradients conjugués.

D'un autre côté, Irwan et Lane ont montré que la décroissance progressive du poids de la régularisation est bénéfique à la reconstruction de phase, car elle permet de lisser des petites irrégularités locales rencontrées au cours de l'optimisation. J'ai vérifié ce bénéfice lors de mes premiers essais, et observé que la vitesse et les amplitudes adaptées de décroissance de μ varient beaucoup en fonction des conditions particulières de turbulence et de bruit. Il m'est donc apparu nécessaire de mettre en oeuvre une stratégie d'ajustement automatique du poids de la régularisation pour pouvoir m'intéresser à des D/r_0 et des niveaux de bruits très différents. C'est pourquoi j'ai finalement opté pour un algorithme plus simple de quasi-Newton de type Levenberg-Marquardt, me permettant de transcrire facilement la méthode d'ajustement automatique de μ développée par Skilling and Bryan [1984], et puisque les fonctions d'adéquation expérimentale et a priori s'écrivent toutes les deux sous la forme de moindres carrés.

3.5.2 Ajustement automatique et à la volée du poids de la régularisation

L'ajustement automatique de μ permet comme on vient de l'expliquer de s'adapter à des conditions de turbulence et de bruit très diverses. Dans tous les cas on souhaite que la régularisation influe suffisamment dans l'optimisation et surtout au début,

lorsque la dépression de f_{post} correspondant au minimum global n'a pas encore été atteinte, et afin de lever les ambiguïtés strictes de périodicité modulo 2π des échantillons de phase. Le poids μ de la régularisation est ainsi ajusté à chaque itération en s'assurant de l'influence suffisante de la régularisation dans le pas. Cela est réalisé par un ajustement de μ tel que la réduction expérimentale soit plus faible avec régularisation que sans :

$$\delta f_{\text{exp}}^{\text{quad}}(\delta \mathbf{p}(\mu)) \simeq \epsilon \times \delta f_{\text{exp}}^{\text{quad}}(\delta \mathbf{p}(\mu = 0)), \quad (3.23)$$

où $\epsilon \in]0, 1[$ est la fraction de la réduction sans régularisation qu'on considère, et où le pas de quasi-Newton $\delta \mathbf{p}(\mu)$ s'écrit en fait

$$\delta \mathbf{p}(\mu) = -(\mathbf{J}^T \cdot \mathbf{J} + \mu \times \mathbf{H}_{\text{prior}})^{-1} \cdot (\mathbf{g}_{\text{exp}} + \mu \times \mathbf{g}_{\text{prior}}). \quad (3.24)$$

En pratique je fais varier μ par dichotomie, comme cela est implémenté par Pichon and Thiébaud [1998] et jusqu'à obtenir l'égalité de l'équation (3.23) à 5% près.

En définitive cette stratégie s'avère particulièrement efficace et automatique, puisqu'elle entraîne une diminution progressive du poids de la régularisation, au fur et à mesure de chaque reconstruction de phase, et puisque la valeur de la fraction ϵ n'a même pas besoin d'être changée avec le D/r_0 . Pour les toutes dernières itérations de l'optimisation, ϵ peut être légèrement accru pour améliorer la convergence, en favorisant le critère expérimental lorsque le niveau de bruit n'est pas trop important.

3.5.3 Projection dans un sous-espace de directions privilégiées

L'algorithme de Levenberg-Marquardt nécessite d'effectuer des calculs longs et de stocker des grosses matrices dans la mémoire du calculateur. En effet, il faut calculer à chaque itération la matrice jacobienne $N \times M$ où N est le nombre de paramètres et M le nombre de données, ce qui peut entre autres nécessiter N transformées de Fourier dans le cas de la reconstruction de phase. Il faut également calculer et stocker l'approximation de la hessienne $\mathbf{J}^T \cdot \mathbf{J}$ de taille $N \times N$. Ainsi la méthode devient vite irréalisable pour des D/r_0 croissants, qui entraînent une augmentation parallèle de N et de M .

J'ai donc également choisi de transposer la variante à mémoire limitée de l'algorithme de Levenberg-Marquardt développée par Skilling and Bryan puis Pichon and Thiébaud [1998], au cas de la reconstruction de phase, afin de pouvoir considérer des D/r_0 plus élevés (plus de degrés de liberté). Dennis and Moré [1974] ont montré que pour que les algorithmes à mémoire limitée convergent rapidement, il est nécessaire que les directions de Newton soient approchées asymptotiquement. Dans cette idée, la méthode de Bryan et Skilling s'emploie à exprimer le pas en fonction de quelques directions privilégiées, directement dérivées des approximations quadratiques de Taylor :

$$\delta \mathbf{p} = \sum_{i=1}^n \beta_i \mathbf{q}_i, \quad (3.25)$$

avec n typiquement de l'ordre de 5, qui est très petit devant le nombre de paramètres. Une fois les directions privilégiées d'optimisation $\mathbf{q}_i(\mathbf{p})$ calculées, le développement de Taylor de $f_{\text{post}}(\mathbf{p})$ est réexprimé dans le sous-espace des nouveaux paramètres $\boldsymbol{\beta}$ sous la forme

$$\delta f^{\text{quad}}(\boldsymbol{\beta}) = \mathbf{g}'_{\text{post}}{}^{\top} \boldsymbol{\beta} + \frac{1}{2} \boldsymbol{\beta}^{\top} \mathbf{A}'_{\text{post}} \boldsymbol{\beta}, \quad (3.26)$$

avec $[\mathbf{A}'_{\text{post}}]_{k,\ell} = \mathbf{q}_k^{\top} \cdot \mathbf{A}_{\text{post}} \cdot \mathbf{q}_{\ell}$ et $[\mathbf{g}'_{\text{post}}]_j = \mathbf{q}_j^{\top} \cdot \mathbf{g}_{\text{post}}$. Les coefficients $\boldsymbol{\beta}$ sont ainsi obtenus de la même façon que les paramètres \mathbf{p} dans l'algorithme de Newton :

$$\boldsymbol{\beta}^+ = -\mathbf{A}'_{\text{post}}{}^{-1} \mathbf{g}'_{\text{post}} \quad (3.27)$$

La première direction est une approximation du pas de quasi-Newton :

$$\mathbf{q}_1(\mathbf{p}) = -\mathbf{B} \cdot \mathbf{g}_{\text{post}}(\mathbf{p}). \quad (3.28)$$

Elle est calculée rapidement à l'aide d'une matrice \mathbf{B} , qui approxime $\mathbf{H}_{\text{post}}^{-1} = (\mathbf{H}_{\text{exp}} + \mu \mathbf{H}_{\text{prior}})^{-1}$, et qui ne dépend pas des paramètres \mathbf{p} . La recommandation de Skilling et Bryan est de prendre

$$\mathbf{B} = \frac{1}{\mu} \mathbf{H}_{\text{prior}}^{-1}. \quad (3.29)$$

En effet cette approximation ne dépend pas des paramètres et des itérations, à l'ajustement de μ près ; elle contraint les paramètres reconstruits avec la bonne métrique (pourvu que les a priori soient bons) ; et $\mathbf{H}_{\text{prior}}$ est positive définie dans le cas d'incertitudes gaussiennes sur les a priori et peut donc être inversée. En reconstruction de phase, cette approximation présente également l'avantage que l'opérateur $\mathbf{H}_{\text{prior}}^{-1} = \mathbf{K}^{\top} \cdot \mathbf{C}_{\varphi} \cdot \mathbf{K}$ est codé de façon creuse, comme on l'a vu à la section 3.3.3.1. Il nécessite donc très peu de mémoire pour être stocké, et peut être rapidement appliqué à un vecteur. Enfin la régularisation étant maintenue à un niveau significatif par rapport au critère expérimental, en particulier au début lorsque l'optimisation est la plus critique, cette approximation est assez fidèle à $\mathbf{H}_{\text{post}}^{-1}$. Notons qu'en pratique, les directions $\mathbf{q}_i(\mathbf{p})$ du sous-espace sont calculées à chaque itération avant l'ajustement de μ , car ce dernier requiert le calcul de \mathbf{A}' et \mathbf{g}' . Ainsi le μ présent dans \mathbf{B} et dans $\mathbf{g}_{\text{post}} = \mathbf{g}_{\text{exp}} + \mathbf{g}_{\text{prior}}(\mu)$ servant à calculer \mathbf{q}_1 , est en fait le μ calculé à l'itération précédente. C'est la raison pour laquelle, je préfère séparer les deux composantes expérimentales et a priori du gradient de l'équation (3.28). Cela permet de laisser une flexibilité de réajustement du poids relatif de la composante a priori, et amène à considérer les deux directions :

$$\mathbf{q}_1 = -\mathbf{B} \cdot \mathbf{g}_{\text{exp}} \quad (3.30)$$

$$\mathbf{q}_2 = -\mathbf{B} \cdot \mathbf{g}_{\text{prior}} \quad (3.31)$$

Les directions suivantes sont déduites en considérant que, une fois \mathbf{q}_i choisie et conduisant à un pas $\delta \mathbf{p}_i = \beta_i \mathbf{q}_i$, la direction qui approxime le pas de quasi-Newton mis à jour

s'écrit

$$\begin{aligned} \mathbf{q}_{i+1} &= -\mathbf{B} \cdot \mathbf{g}(\mathbf{p} + \delta \mathbf{p}_i) \\ &\simeq -\mathbf{B} \cdot [\mathbf{g}(\mathbf{p}) + \mathbf{H}(\mathbf{p}) \cdot \delta \mathbf{p}_i] \\ &\propto \mathbf{q}_i - \text{cste} \times \mathbf{B} \cdot \mathbf{H} \cdot \mathbf{q}_i . \end{aligned} \quad (3.32)$$

Pour la même raison que précédemment relative à la mise à jour de μ , on sépare les composantes expérimentales et a priori de \mathbf{H} . Comme $\mathbf{B} \propto \mathbf{H}_{\text{prior}}^{-1}$, la direction issue de la composante a priori de \mathbf{H} et de \mathbf{q}_2 est proportionnelle à \mathbf{q}_2 et ne présente pas d'intérêt. En pratique je prends donc

$$\mathbf{q}_3 = -\mathbf{B} \cdot \mathbf{H}_{\text{exp}} \cdot \mathbf{q}_1 \quad (3.33)$$

$$\mathbf{q}_4 = -\mathbf{B} \cdot \mathbf{H}_{\text{exp}} \cdot \mathbf{q}_3 . \quad (3.34)$$

Finalement pour réajuster le pas effectué à l'itération $(n-1)$ et s'assurer que le nouveau pas à l'itération (n) n'annule pas exactement l'ancien, j'ai choisi explicitement la direction

$$\mathbf{q}_5^{(n)} = \delta \mathbf{p}^{(n-1)} . \quad (3.35)$$

J'ai observé en pratique que ce choix accélère la convergence de l'algorithme.

3.5.4 Région de confiance

L'importance du contrôle de la longueur des pas de Newton a été discutée à la section 2.5.5. J'ai choisi d'utiliser l'algorithme de région de confiance présenté à cette même section pour ajuster la longueur du pas en fonction de l'accord quadratique. La reparamétrisation dans le sous-espace implique cependant quelques ajustements dans l'algorithme initial de région de confiance de Moré and Sorensen [1983]. En effet pour conserver la métrique a priori des paramètres, il faut imposer la norme suivante aux sous-paramètres $\boldsymbol{\beta}$:

$$\|\boldsymbol{\beta}\|_{\mathbf{M}}^2 = \frac{1}{2} \boldsymbol{\beta}^\top \cdot \mathbf{M} \cdot \boldsymbol{\beta} = \|\delta \boldsymbol{\phi}\|^2 , \quad (3.36)$$

avec $[\mathbf{M}]_{k,l} = \mathbf{q}_k^\top \cdot \mathbf{q}_l$. Le pas de région de confiance s'écrit ainsi

$$\begin{aligned} \delta \mathbf{p}^{\text{TR}} &= \boldsymbol{\beta}^{\text{TR}}(\mu, \ell) \cdot \mathbf{M} \\ &= - \left[\left(\mathbf{A}'_{\text{post}}(\mu) + \ell \mathbf{M} \right)^{-1} \cdot \mathbf{g}'_{\text{post}}(\mu) \right] \cdot \mathbf{M} , \end{aligned} \quad (3.37)$$

où ℓ est ajusté par l'algorithme de Moré et Sorensen.

3.5.5 Calcul rapide des dérivées partielles

Dans le cas où les nombres de paramètres N et de données M sont élevés, il est nécessaire de pouvoir appliquer le jacobien expérimental \mathbf{J}_{exp} sans le stocker et de façon efficace pour que l'algorithme soit rapide. En effet l'algorithme d'optimisation requiert le calcul de

$$\mathbf{g}_{\text{exp}} = \mathbf{J}_{\text{exp}}^{\top} \cdot \mathbf{W}_{\text{exp}} \cdot (\mathbf{d} - \mathbf{m}) \quad (3.38)$$

$$[\mathbf{A}'_{\text{exp}}]_{k,l} = (\mathbf{J}_{\text{exp}} \cdot \mathbf{q}_k)^{\top} \cdot \mathbf{W}_{\text{exp}} \cdot (\mathbf{J}_{\text{exp}} \cdot \mathbf{q}_l) \quad (3.39)$$

$$\mathbf{q}_{3,4} = -\mu \times \mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{J}_{\text{exp}}^{\top} \cdot \mathbf{J}_{\text{exp}} \cdot \mathbf{q}_{1,3}, \quad (3.40)$$

à chaque itération. L'expression directe du jacobien dans le cas de la reconstruction de phase s'écrit (cf. section 3.2.1 pour les notations)

$$[\mathbf{J}_{\text{exp}}]_{i,j} \triangleq -\frac{\partial m_i(\mathbf{p})}{\partial p_j} = 2\alpha \operatorname{Im} \left(\hat{a}_i^* \sum_k F_{i,k} a_k \frac{\partial \varphi_k}{\partial p_j} \right) \quad (3.41)$$

En réordonnant les opérateurs pour le calcul rapide on obtient les expressions suivantes pour l'application de l'opérateur jacobien et de son transposé à un vecteur :

$$(\mathbf{J}_{\text{exp}} \cdot \mathbf{q})_j = 2\alpha \operatorname{Im} \left(\hat{a}_j^* \times \sum_k F_{j,k} a_k \left(\sum_l \frac{\partial \varphi_k}{\partial p_l} q_l \right) \right) \quad (3.42)$$

$$(\mathbf{J}_{\text{exp}}^{\top} \cdot \mathbf{q})_l = 2\alpha \sum_k \frac{\partial \varphi_k}{\partial p_l} \times \operatorname{Im} \left(a_k \sum_j F_{k,j} \hat{a}_j^* q_j \right). \quad (3.43)$$

Ces calculs ne nécessitent en effet qu'une seule transformée de Fourier.

3.6 Simulation de données

Je présente ici comment les données servant par la suite pour la caractérisation ont été simulées, et je résume dans un tableau les différentes conditions de bruit et de turbulence envisagées.

Les écrans de phase considérés suivent une statistique spatiale de Kolmogorov. La méthode la plus simple pour les simuler consiste à calculer la matrice de covariance spatiale de Kolmogorov \mathbf{C}_{φ} et à en faire la décomposition en valeurs singulières, comme cela est décrit à l'équation (3.22). La statistique spatiale des phases est en effet obtenue en appliquant la matrice $\mathbf{U} \cdot \sqrt{\Sigma}$ à un vecteur aléatoire de la loi normale $\mathcal{N}(0, 1)$. Cependant la matrice \mathbf{C}_{φ} et sa décomposition en valeurs singulières, deviennent rapidement trop volumineuses avec l'augmentation du D/r_0 , pour le stockage et le calcul. Pour effectuer la simulation des phases plus rapidement et sans utiliser de mémoire,

j'ai donc utilisé un générateur fractal, décrit dans [Béchet et al., 2006] et semblable à celui de Lane et al. [1992]. La différence sur la statistique des phases simulées par rapport au calcul explicite de C_φ est très négligeable (ce qui est caractérisé dans les articles en question).

Par ailleurs j'ai choisi de toujours échantillonner les images à ~ 2 pixels CCD par λ/D , pour d'une part échantillonner les tavelures et ainsi conserver toute l'information haute fréquence, et d'autre part limiter la dilution du signal sur le détecteur afin de maximiser le rapport signal à bruit. En pratique cet échantillonnage est réalisé en plongeant la pupille du télescope simulée dans un support de taille double avant de lui appliquer la transformée de Fourier discrète pour le calcul de l'image.

Le bruit de photons est simulé dans les images à l'aide d'un générateur pseudo-aléatoire suivant la statistique de Poisson, tandis qu'un générateur pseudo-aléatoire gaussien est utilisé pour ajouter au signal un bruit de détecteur assez faible. Ce-dernier est choisi spatialement homogène avec $\sigma_{\text{CCD}}=0.32$ photons/pixel pour chaque pixel de l'image. Le bruit de photons domine ainsi le bruit de détecteur, pour l'ensemble des simulations.

L'augmentation du nombre de tavelures avec le D/r_0 entraîne une dilution du signal, à nombre égal de photons dans les images. J'ai donc également calculé un rapport signal à bruit par speckle pour chaque couple de turbulence et de bruit envisagé pour la reconstruction de phase, afin de quantifier le rapport signal à bruit sur l'information haute fréquence de l'image :

$$\text{RSB} = \frac{N_{\text{ph}}/N_{\text{sp}}}{\sqrt{N_{\text{ph}}/N_{\text{sp}} + \sigma_{\text{CCD}}^2 \times s_{\text{sp}}^2}}, \quad (3.44)$$

où N_{ph} est le nombre moyen de photons dans l'image, $N_{\text{sp}} \sim (D/r_0)^2$ est le nombre de speckles, et s_{sp} est la taille moyenne d'un speckle. L'ensemble des conditions envisagées est résumé dans le tableau ci-dessous :

TAB. 3.1 – Paramètres envisagés dans les différentes simulations

Id.	N_{ph}	D/r_0	RSB	Id.	N_{ph}	D/r_0	RSB
1	50 000	5	44.7	6	50 000	8	27.9
2	50 000	7	31.9	7	22 000	8	18.5
3	50 000	9	24.8	8	9 000	8	11.8
4	50 000	10	22.4	9	4 000	8	7.9
5	50 000	11	18.1	10	1 500	8	4.8

3.7 Optimisation globale de phases décorréliées

3.7.1 Stratégie pseudo-aléatoire

Le cas d'optimisation globale le plus difficile à traiter, est celui dans lequel les phases à reconstruire ne possèdent aucune corrélation temporelle entre elles. En effet, comme on ne peut pas s'appuyer sur la reconstruction de phase précédente pour initialiser chaque optimisation d'une nouvelle phase avec une estimation grossière, il faut pouvoir résoudre toutes les ambiguïtés strictes et non-strictes des critères d'adéquation $f_{\text{exp}}(\mathbf{p})$ et $f_{\text{post}}(\mathbf{p})$. En pratique j'ai vérifié que les ambiguïtés deviennent trop nombreuses au-delà de $D/r_0=4$, pour pouvoir estimer correctement toutes les phases en utilisant uniquement l'algorithme d'optimisation locale que je viens de décrire. Il faut alors disposer d'une stratégie d'optimisation globale, dotant l'algorithme local d'une initialisation de phase suffisamment proche de la phase à estimer.

En l'absence d'alternative plus efficace, la stratégie la plus rudimentaire pour obtenir cette initialisation, mais néanmoins constructive, consiste à effectuer des tirages aléatoires de phase, jusqu'à obtenir une phase pour laquelle l'adéquation a posteriori $f_{\text{post}}(\mathbf{p})$ est suffisamment bonne. En effet, de la même façon que j'ai jusqu'ici supposé que le minimum global de $f_{\text{post}}(\mathbf{p})$ correspond toujours à une très bonne estimation de la phase — à l'ambiguïté entre $\varphi^+(\mathbf{r})$ et $-\varphi^+(-\mathbf{r})$ près —, il semble tout aussi raisonnable en s'appuyant sur des arguments de continuité, de supposer que dans une certaine zone de voisinage autour du minimum global, les faibles valeurs de $f_{\text{post}}(\mathbf{p})$ correspondent aussi à des bonnes estimations de la phase. Le nombre de paramètres à estimer pour pouvoir obtenir une adéquation correcte avec l'image étant relativement important, de l'ordre de $(D/r_0)^2$, cette stratégie aléatoire semble toutefois fort coûteuse à première vue, et même en utilisant des méthodes de type Monte-Carlo pour réduire le nombre de tentatives. Toutefois les observations que j'ai réalisées lors de mes simulations m'ont permis de raffiner cette stratégie.

J'ai tout d'abord observé que la zone de dépression de l'adéquation $f_{\text{post}}(\mathbf{p})$ autour de l'optimum global est plus large que les zones de dépressions autour des minima locaux. De plus, le raffinement d'une zone de dépression vers le minimum local, lorsque la zone de dépression est suffisamment large pour que l'algorithme d'optimisation locale s'y arrête, consitue la partie la plus longue de l'optimisation, notamment parce que le contrôle du pas par la région de confiance y prend une plus grande importance pour assurer la convergence.

Ainsi il apparaît judicieux de combiner la stratégie des tirages aléatoires à quelques itérations de l'algorithme local, c'est à dire suffisamment peu pour ne pas entamer l'exploration détaillée d'un minimum local. En pratique, j'ai constaté que de l'ordre de 60 itérations sont suffisantes pour pouvoir estimer d'après la valeur de $f_{\text{exp}}(\mathbf{p})$ si la zone de dépression du minimum global a été atteinte. Ce nombre d'itérations dépend légèrement du D/r_0 et des conditions de bruit et peut être optimisé conjointement avec

le nombre de départs aléatoires pour accélérer l'algorithme global.

Les départs aléatoires que j'ai considérés, sont des tirages de la distribution de matrice de covariance $\mathbf{U}\Sigma^{1/2}$, qui est obtenue grâce à la décomposition en valeur singulière de \mathbf{C}_φ décrite à l'équation (3.22). Ils constituent donc un échantillonnage des modes singuliers de la turbulence. En pratique je n'ai pas constaté de variation significative des performances, à diminuer le nombre de modes singuliers soumis au tirage aléatoire, pourvu toutefois que ce nombre reste suffisamment élevé, c'est à dire typiquement $\geq 2D/r_0$.

3.7.2 Corrélation entre f_{exp} et l'erreur de reconstruction de la phase

La corrélation entre les valeurs de f_{exp} et les erreurs de reconstruction de la phase, qu'on obtient à l'issue de la stratégie d'optimisation globale que je viens de décrire, est représentée sur la figure 3.2 pour différentes conditions de turbulence et de bruit. Pour chaque couple de D/r_0 et de rapport signal à bruit, 250 phases différentes ont été considérées, en effectuant pour chacune un nombre délibérément excessif de 200 départs aléatoires.

L'ensemble de ces graphiques montre que les meilleures reconstructions de phase, correspondent comme on l'espérait aux plus faibles valeurs du critère d'adéquation expérimental f_{exp} . Cela justifie aussi l'utilisation d'un seuil de f_{exp} comme critère d'arrêt de la stratégie d'optimisation globale. Ce seuil peut être déterminé sur les graphiques en fonction du bruit et de D/r_0 , comme la valeur de l'adéquation pour laquelle les erreurs de reconstruction sont toujours inférieures à 1 radian. J'ai en effet vérifié que le minimum global est toujours atteint en raffinant la phase, si celle-ci est estimée avec cette précision de 1 radian. Pour le traitement de vraies données en temps-réel, l'estimation de D/r_0 et du niveau de bruit peut ne pas être directement accessible. Il suffit alors par exemple de procéder par auto-calibration, en effectuant un nombre suffisant de reconstructions à l'initialisation et en ajustant le seuil au fur et à mesure des reconstructions et avec l'évolution temporelle de r_0 .

La comparaison des deux graphiques du haut montre la nette augmentation du nombre de mauvaises estimations de la phase entre $D/r_0=5$ et $D/r_0=11$, et traduit l'augmentation du nombre de minima locaux avec le D/r_0 . La distribution de f_{exp} pour $D/r_0=11$ devient bimodale, avec la zone des minima globaux de f_{exp} , pour laquelle les valeurs du critère sont faibles et les phases bien reconstruites, et la zone des minima locaux pour laquelle les valeurs du critère sont nettement moins bonnes tout comme l'erreur de reconstruction de phase.

La comparaison des deux graphiques du bas montre la dégradation de la corrélation entre les faibles valeurs de f_{exp} et les bonnes estimations de la phase, lorsque le rapport signal à bruit diminue. En effet les deux modes de la distribution ont tendance à se rapprocher à faible flux et on conçoit qu'il arrive un moment où la stratégie d'optimisation globale ne peut plus fonctionner, si on diminue encore le nombre de photons.

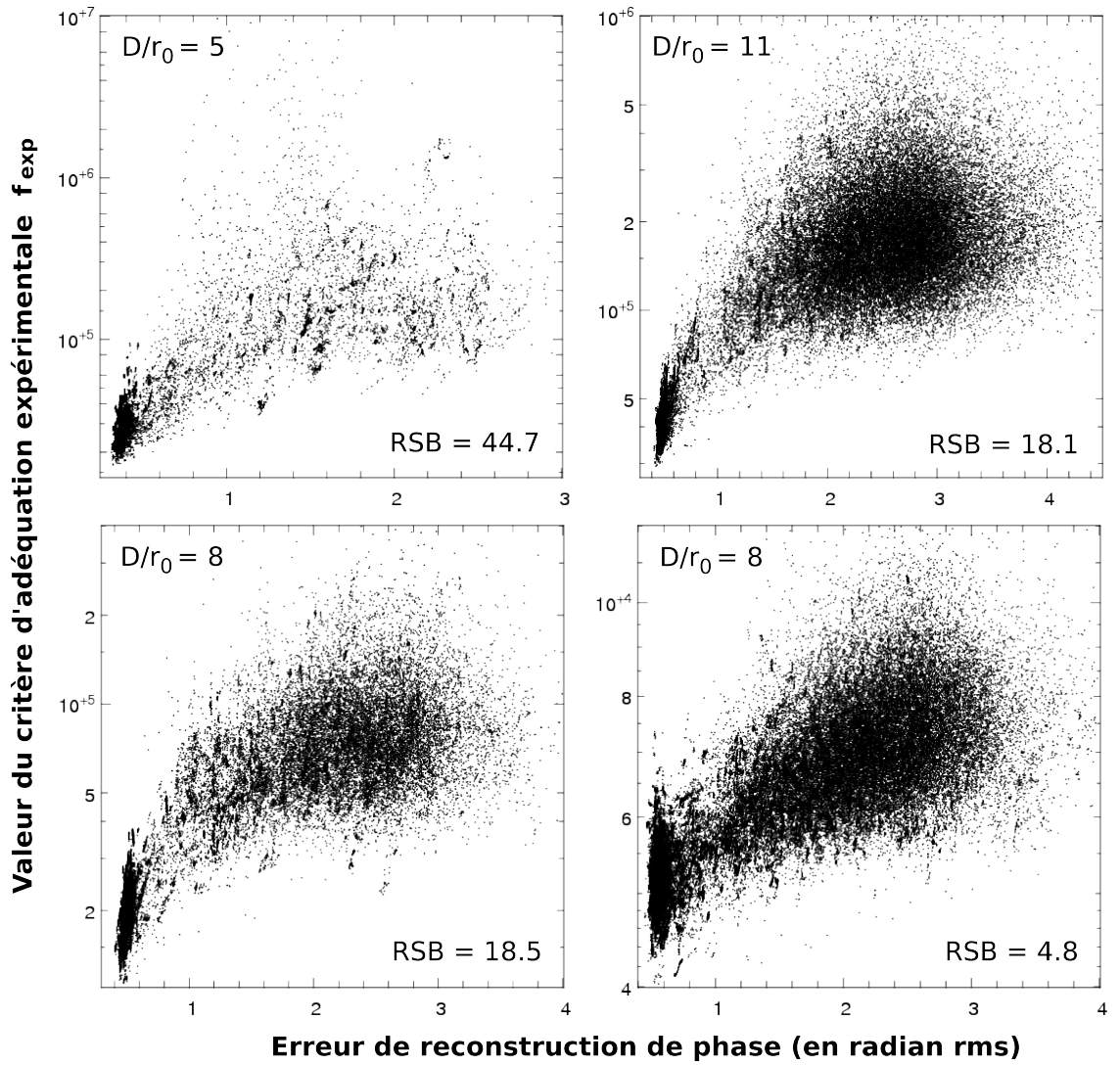


FIG. 3.2 – Valeurs de f_{exp} en fonction de la déviation standard de l'erreur de reconstruction de la phase (en radian) et pour différentes conditions de turbulence (D/r_0) et de bruit dans l'image. Pour chaque simulation, 250 différentes phases sont considérées, et 200 départs aléatoires sont testés pour chaque phase. Les points sont obtenus à l'issue de 50 itérations de l'algorithme local pour chaque départ aléatoire.

3.7.3 Efficacité de la stratégie en fonction du D/r_0

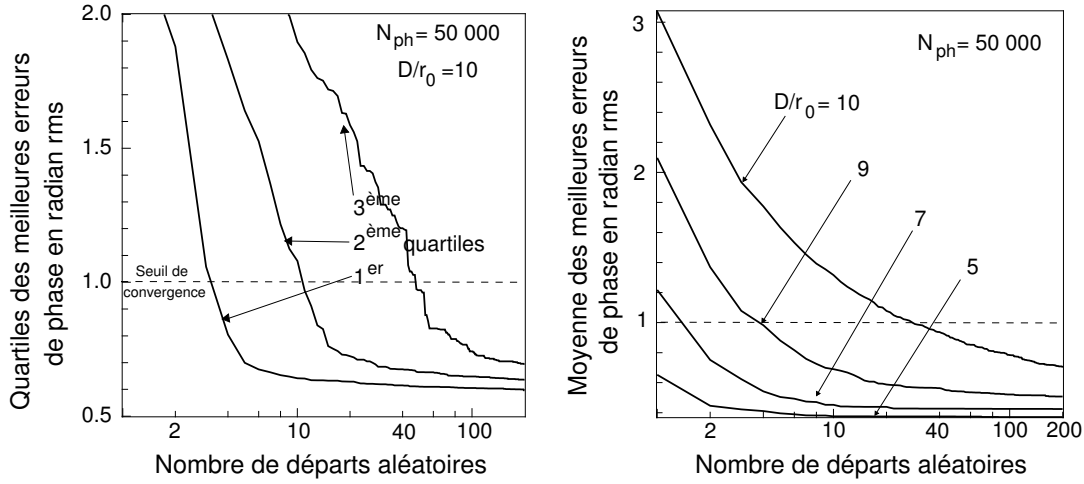


FIG. 3.3 – Caractérisation statistique du nombre de départs aléatoires nécessaires pour atteindre une certaine erreur de reconstruction de la phase, en fonction du D/r_0 .

Comme je me suis assuré qu'il suffit d'atteindre un certain seuil $f_{\text{exp}}^{\text{seuil}}$ pour converger à coup sûr vers le minimum global, l'efficacité de la stratégie pseudo-aléatoire se mesure par le nombre moyen de départs aléatoires nécessaires pour atteindre ce seuil. J'ai donc représenté sur la figure 3.3 les interquartiles et la moyenne de la meilleure erreur de phase obtenue en fonction du nombre de départs aléatoires. Pour chaque couple de turbulence et de rapport signal à bruit, le moyennage concerne au moins 250 phases. On observe l'augmentation du nombre de départs aléatoires avec la force relative de la turbulence D/r_0 , qui traduit un nombre croissant de minima locaux. Cette augmentation est aussi renforcée pour $D/r_0=10$ par un mauvais échantillonnage du modèle. En effet celui-ci n'est pas assez fin pour que les a priori et la contrainte de lissage apportée par la reparamétrisation, puisse imposer l'absence de saut de période entre deux échantillons, et ainsi lever toutes les ambiguïtés strictes de pliage modulo 2π de la phase, comme cela est discuté plus en détail à la section 3.9.3.

3.7.4 Efficacité de la stratégie en fonction du bruit

De la même façon que pour la force de la turbulence, j'ai représenté la dépendance statistique du nombre de départs aléatoires nécessaire pour atteindre une précision donnée de reconstruction de la phase, en fonction du niveau de bruit et en considérant un D/r_0 fixe de 8. On voit qu'il n'y a pas de dégradation continue des performances de la stratégie avec le bruit, mais une brusque dégradation pour le nombre de photons le plus faible. En regardant plus en détail sur les ailes de la distribution (cas extrêmes au-delà

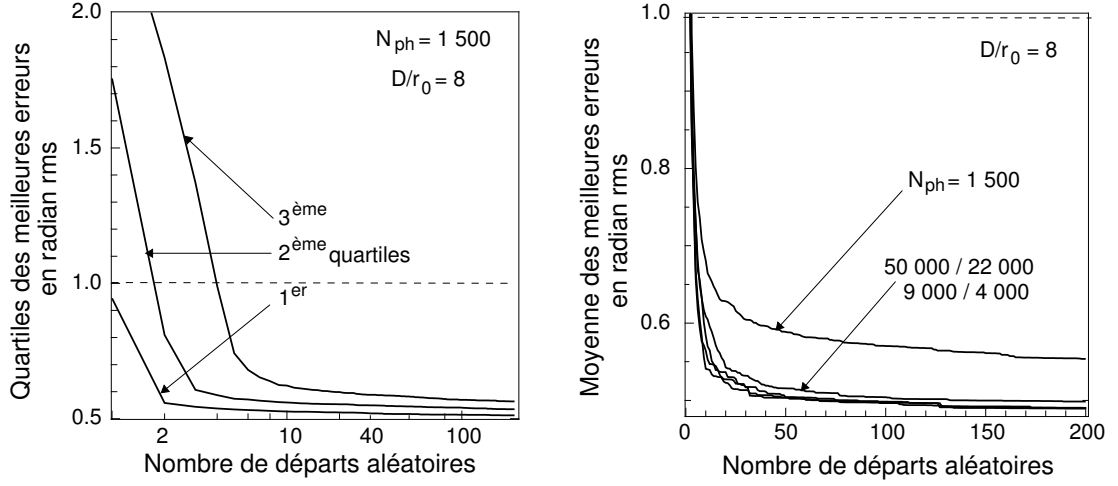


FIG. 3.4 – Caractérisation statistique du nombre de départs aléatoires nécessaire pour atteindre une certaine erreur de reconstruction de la phase, en fonction du nombre moyen de photons dans les images.

du troisième interquartile) il s'avère que certains cas de reconstruction commencent à ne plus converger. J'ai pu toutefois vérifier que l'augmentation du nombre de départs aléatoires permet de gérer ces cas les plus difficiles.

3.7.5 Convergence finale de l'algorithme

J'ai représenté sur la figure 3.5 la précision finale de reconstruction de la phase, en fonction des D/r_0 et des niveaux de bruits précédemment considérés. Cette précision est obtenue en raffinant la phase qui correspond au meilleur f_{exp} obtenu après 200 départs aléatoires. Comme on l'a vu précédemment, la stratégie la mieux adaptée ne consiste pas à utiliser un nombre fixe de départs aléatoires, mais à tirer un nombre de départs, propre à chaque phase, pour atteindre une valeur seuil de f_{exp} . Le nombre limité de départs aléatoires explique donc ici que $\sim 15\%$ des phases ne sont pas bien estimées pour $D/r_0=8$ et $N_{ph}=1500$ photons (ce que j'ai vérifié).

On observe sur la figure de gauche, une légère dégradation de la reconstruction avec le D/r_0 . Celle-ci est liée à l'augmentation du bruit de speckle dans l'image avec le D/r_0 à flux constant (cf. RSB dans la table 3.1). Sur la figure de droite on constate la dégradation de l'erreur de reconstruction avec la diminution du nombre de photons dans l'image. Cette dégradation évolue à première vue comme la racine carrée du nombre de photons dans l'image. La précision que j'obtiens n'est sans doute pas la meilleure possible à cause de la présence de petits minima locaux, en nombre variable avec le D/r_0 et avec le niveau de bruit, et une optimisation de l'algorithme dans sa phase finale combinée à une étude analytique me semble donc devoir encore être menée pour pouvoir tirer des conclusions précises à propos des performances optimales de

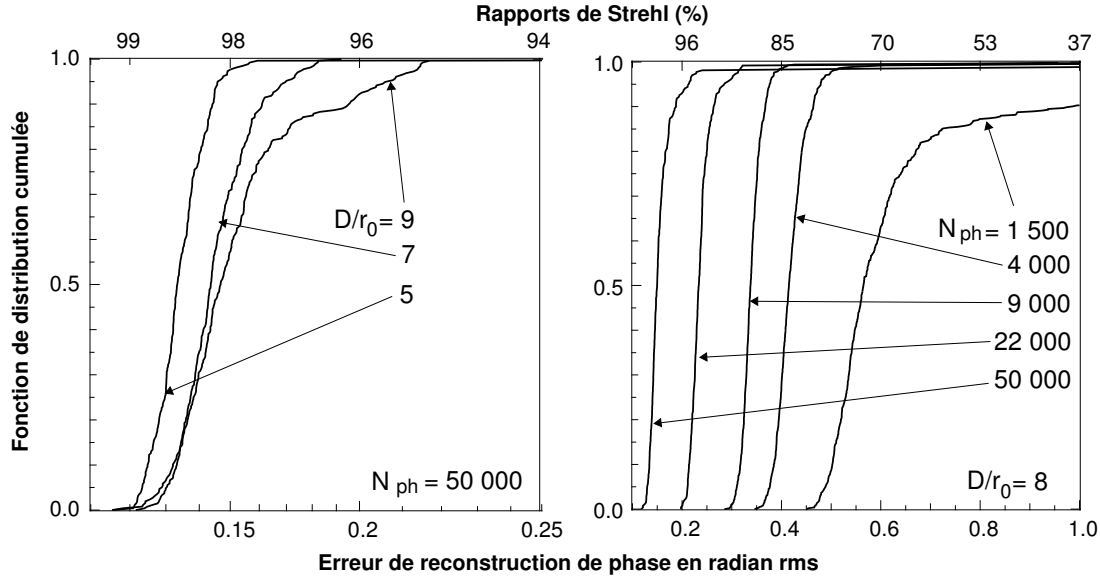


FIG. 3.5 – Représentation statistique de la précision finale de reconstruction de phase en fonction du D/r_0 et du niveau de bruit

convergence. Globalement, les résultats obtenus montrent cependant que la phase peut-être reconstruite très précisément et jusqu’à un nombre de photon relativement faible.

Le cas $D/r_0 = 11$ a par ailleurs fait l’objet d’une étude particulière en fonction de l’échantillonnage, et est détaillé à part à la section 3.9.3.

3.8 Optimisation globale pour des phases à forte corrélation temporelle

Dans le cas d’acquisitions continues, qui se présenterait par exemple en couplant en temps-réel la reconstruction de phase avec une Optique Adaptative, on peut penser à exploiter la corrélation temporelle des phases dans l’optimisation, pourvu que les trames soient suffisamment rapprochées par rapport à la vitesse caractéristique d’évolution de la turbulence. On souhaite en effet ainsi ne plus avoir besoin de la stratégie d’optimisation globale — hormis pour la première phase et pour pouvoir gérer les éventuelles instabilités dans la corrélation temporelle — et n’avoir ainsi simplement qu’à raffiner la phase reconstruite à l’itération précédente avec l’algorithme local, pour chaque nouvelle image.

J’ai testé cette stratégie temporelle en fonction du niveau de corrélation entre les phases successives pour un $D/r_0=11$ et un niveau de bruit de $N_{ph}=10^5$ photons. Pour les simulations j’ai choisi de simplement quantifier la vitesse d’évolution de la tur-

bulence entre deux trames, par la translation $\Delta \mathbf{r}$ d'une seule couche turbulente, en utilisant l'hypothèse de Taylor pour l'évolution temporelle de la turbulence. En réalité, la translation d'une seule couche turbulente ne permet pas de simuler le bouillonnement des tavelures qu'on observe sur de vraies données [Roddier et al., 1982]. Cependant j'ai vérifié en simulant la translation de plusieurs couches, que la stratégie n'est pas affectée par le nombre de couches, et que son efficacité ne dépend que de la précision rms de l'initialisation de la phase.

Comme on le voit sur la figure 3.6, représentant la distribution cumulative des erreurs de reconstruction de phase lorsque les phases se translatent de $\Delta \mathbf{r} = \mathbf{r}_0$ entre deux trames et lorsqu'on utilise seulement l'algorithme d'optimisation local combiné à la stratégie temporelle, on arrive bien à reconstruire toutes les phases sans exception avec cette stratégie. De plus, comme on s'y attendait, l'ambiguïté entre $\varphi(\mathbf{r})$ et $-\varphi(-\mathbf{r})$ est levée par la continuité temporelle pourvu que la première phase soit la bonne. Le choix de cette première phase peut par exemple se faire en testant laquelle des deux corrections permet d'obtenir le meilleur rapport de Strehl. Autrement on peut aussi imaginer se baser sur l'évolution de plusieurs phases reconstruites pour choisir le bon signe.

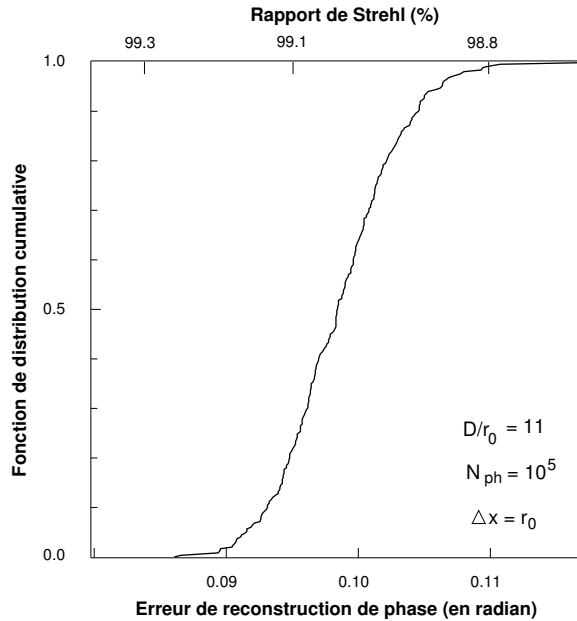


FIG. 3.6 – Représentation statistique de la précision finale de reconstruction de phase obtenue avec la stratégie d'optimisation temporelle. Une dizaine de phase ont fait l'objet d'une cinquantaine de translations chacune.

On pourrait combiner la stratégie d'optimisation globale et la stratégie temporelle pour gérer des conditions plus extrêmes d'évolution de la turbulence. Par ailleurs on

s'attend à pouvoir gérer des évolutions plus fortes grâce à une meilleure contrainte du tip-tilt à partir des images.

3.9 Qualification de l'importance de certains choix

Je qualifie dans cette section l'importance de certains ingrédients dans le bon fonctionnement de l'optimisation globale, en les faisant varier isolément. Cela permet en effet de mettre plus en relief la combinaison complexe de nombreux choix qui a été nécessaire pour résoudre ce problème difficile, et que j'ai effectuée en partie à l'aide de nombreuses simulations et par améliorations progressives ; et cela permet également de quantifier la sensibilité de l'algorithme par rapport à ces choix. Pour caractériser ces ingrédients j'ai considéré le cas le plus difficile de la reconstruction de phases décorrélées, car l'importance de chaque ingrédient se quantifie plus probablement dans ce contexte par le nombre statistique d'échecs de la reconstruction, que lorsque l'optimisation est fortement facilitée par la continuité temporelle. Je représente donc dans cette section l'erreur de reconstruction statistique (3^{ème} interquartile) en fonction du nombre de départs aléatoires et pour le cas le plus difficile envisagé de $D/r_0=11$.

3.9.1 Bornage du tip-tilt

L'importance primordiale de la contrainte du tip-tilt a été précédemment discutée. Elle est due comme on l'a vu aux ambiguïtés introduites par la seule translation de la figure de speckles reconstruite, par rapport à celle des données. On voit sur la figure 3.7 que le nombre de départs nécessaire pour reconstruire les phases est 4.4 fois moins important quand le tip-tilt est parfaitement connu et borné tout au long de l'ajustement que quand il est totalement inconnu. Le gain apporté par le barycentre est plus modeste pour $D/r_0=11$, car le barycentre et le tip-tilt deviennent des grandeurs assez différentes pour de tels D/r_0 . Toutefois ce dernier gain n'est pas négligeable ($\sim 25\%$)

3.9.2 Régularisation de la phase

La régularisation de la phase est essentielle pour lever les dégénérescences de la reconstruction de phase et favoriser l'optimisation globale en gommant certains minima locaux de f_{exp} . D'un point de vue statistique (bayésien), la régularisation de Kolmogorov est la mieux adaptée. Toutefois j'ai comparé son efficacité avec le simple a priori de lissage décrit à la section 3.3.3.2, qui semble à première vue également bien adapté aux nombreuses ambiguïtés de pliage de la phase. Comme on peut l'observer sur la figure 3.8 le nombre de départs est significativement meilleur. Le choix de la bonne régularisation n'est donc pas si anodin que cela : les ambiguïtés sont plus nombreuses

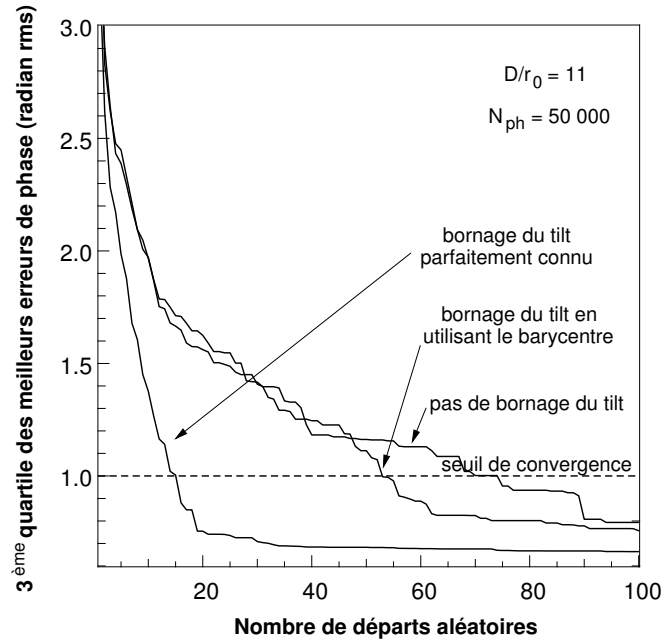


FIG. 3.7 – Importance du bornage du tip-tilt pour la reconstruction de phase.

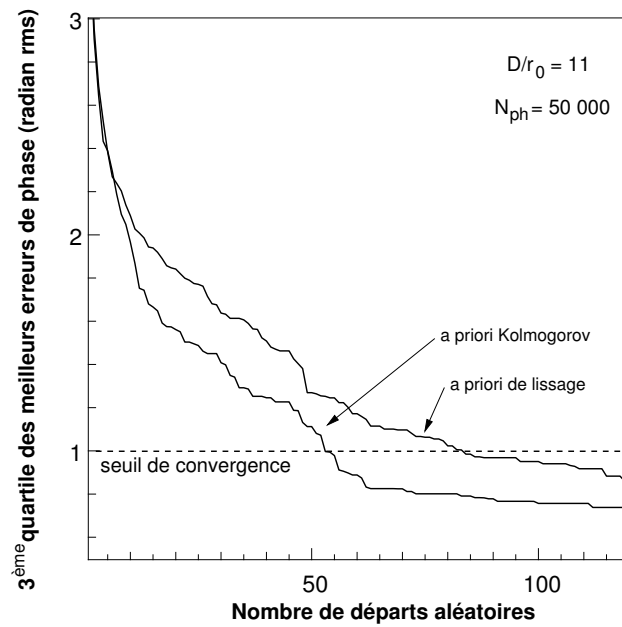


FIG. 3.8 – Importance du choix de la régularisation pour la reconstruction de phase.

que le simple dépliage de la phase et sont mieux traitées par le choix de la meilleure régularisation statistique.

3.9.3 Echantillonnage de la phase

L'échantillonnage de la phase est un ingrédient clé de la reconstruction, à cause des ambiguïtés strictes de pliage de la phase modulo 2π . En effet il faut que la phase soit échantillonnée suffisamment finement dans le modèle pour que le lissage apporté par la régularisation et par le rééchantillonneur de la reparamétrisation puisse garantir l'absence de saut de phase entre deux échantillons, et donc l'absence de cette ambiguïté dans le critère expérimental. J'ai testé pour illustrer cela deux échantillonnages différents de la phase dans le modèle, avec le même D/r_0 et le même échantillonnage des tavelures dans l'image (1 pixel CCD $\approx 0.5\lambda/D$). En pratique cela veut dire que le nombre de pixels de l'image a été varié parallèlement au nombre de pixels de la phase. Comme on le voit sur la figure 3.9, le nombre de départs aléatoires nécessaire est jusqu'à trois fois moins important avec l'échantillonnage le plus fin de 65×65 pixels, lequel permet de mieux démanteler les dégénérescences de pliage de la phase que l'échantillonnage plus grossier. J'ai observé plus généralement, que le diamètre de la pupille doit comprendre au moins $4 \times D/r_0$ pixel, pour assurer une contrainte de lissage suffisante. La précision finale est également représentée sur la figure de droite pour le meilleur échantillonnage, avec la convergence de toutes les phases (à 1% près) et pour $D/r_0=11$.

3.9.4 Région de confiance

L'utilisation d'une région de confiance dans l'algorithme d'optimisation locale est justifiée par la nécessité d'ajuster la longueur des pas de quasi-Newton lorsque le conditionnement de la hessienne est mauvais. Plus généralement elle est basée sur le bon accord entre les variations effectives de la fonction de pénalisation avec sa description quadratique. En effet comme le pas est directement dérivé de cette description quadratique, il semble légitime de baser l'acceptation du pas en fonction de ce critère.

En pratique j'ai observé que la convergence de l'algorithme global est meilleure en débridant la région de confiance (i.e. les bornes de rejet de l'accord quadratique) lors de la première phase d'exploration, c'est à dire lorsque l'algorithme, dans sa phase globale, est encore à la recherche de la zone de dépression du bon minimum. Cela permet en effet de ne pas trop raffiner de petites irrégularités insignifiantes du critère et est justifié par le fait que le volume de dépression du minimum global est plus important que celui des minima locaux. Le resserrement de la région de confiance permet au contraire dans un deuxième temps de converger plus efficacement vers le minimum.

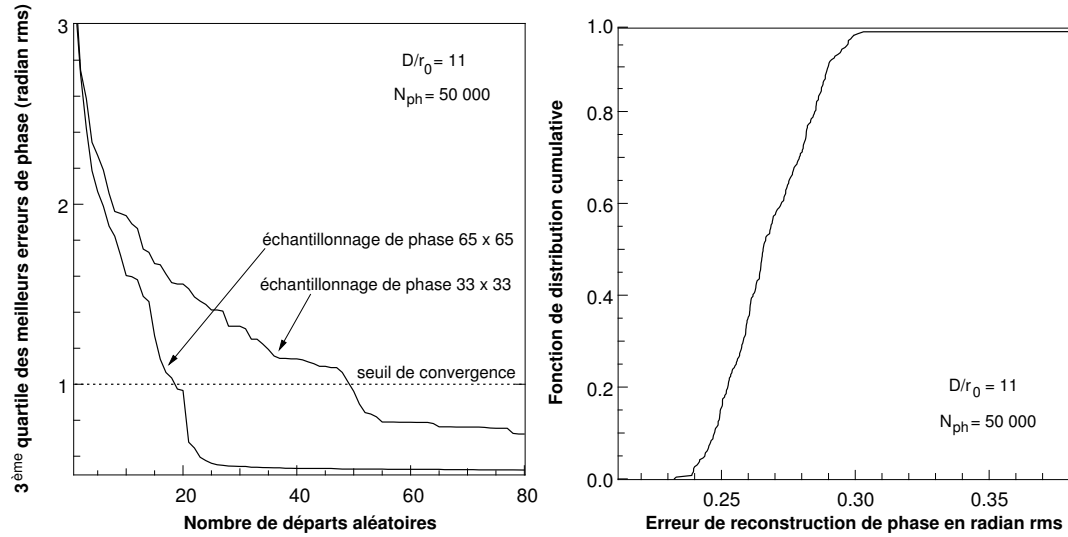


FIG. 3.9 – Importance du niveau d'échantillonnage de la phase dans le modèle pour la reconstruction de la phase dépliée. A gauche on voit le nombre de départs aléatoires nécessaire pour atteindre une certaine erreur de reconstruction de la phase. Tandis que l'erreur de reconstruction finale en raffinant l'estimation obtenue de la première étape de l'optimisation globale est représentée à droite pour le meilleur des deux échantillonnages.

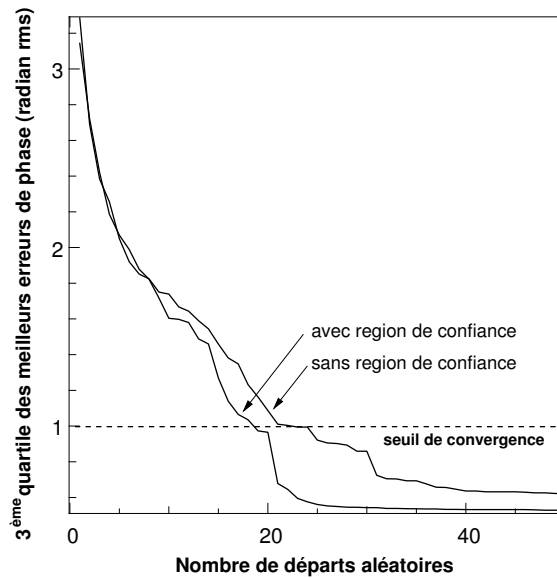


FIG. 3.10 – Influence du contrôle par la région de confiance de la longueur des pas d'optimisation.

Comme on le voit sur la courbe 3.10, il est quand même légèrement préférable de conserver la région de confiance dans les premiers temps de l'optimisation. Cela permet en effet la gestion de très mauvais conditionnements de hessiennes pour accrocher la dépression du minimum global et ne pas en ressortir parce qu'un pas vraiment trop large par rapport à la zone de validité des approximations quadratiques serait effectué à cet endroit.

3.10 Conclusions

J'ai développé dans ce chapitre un algorithme permettant de reconstruire la phase dépliée sur la pupille d'un télescope — d'une onde EM ayant traversé l'atmosphère turbulente — à partir d'une seule image tavelée observée au foyer. C'est un problème inverse non-linéaire nécessitant d'effectuer de l'optimisation globale. J'ai démontré la validité de l'optimisation du critère de maximum de vraisemblance a posteriori pour estimer précisément la phase pour des $D/r_0 > 4$, en développant des stratégies d'optimisation globale. Ces stratégies me permettent aussi bien d'exploiter la corrélation temporelle entre des phases successives, que de reconstruire correctement la phase lorsque la corrélation temporelle est insuffisante pour simplifier la tâche de l'optimisation.

Les résultats que j'ai obtenus à partir de données consciencieusement simulées, concernent des conditions de turbulence beaucoup plus difficiles que ce qui a été réalisé jusqu'à présent avec une seule image et en dépliant la phase complètement. En effet mon algorithme permet aujourd'hui de gérer typiquement 120 degrés de liberté pour un système d'optique adaptative voire plus en exploitant les corrélations temporelles des phases successives, contre 16 obtenus en optimisation locale par [Irwan and Lane, 1998]. Ces résultats ont fait l'objet de la publication dans Josa-A qui est en annexe du manuscrit, et ils constituent par ailleurs une contribution substantielle à la résolution de ce problème, étant donnée la forte augmentation du nombre de minima locaux et de dégénérescences avec la croissance de D/r_0 .

Les limitations actuelles de l'algorithme pour des $D/r_0 \sim 13$ sont dues principalement à la mauvaise estimation du tip-tilt qui est effectuée à partir du barycentre de l'image. Un estimateur du tip-tilt utilisant la distribution d'intensité de l'image me semble donc devoir être développé pour pouvoir dépasser ces limitations, à moins de pouvoir de façon alternative borner efficacement le barycentre dans la phase reconstruite, en se basant sur le gradient moyen de phase discret au lieu de la pente de la phase.

Au cours de ma thèse, j'ai pu tester d'autres approches pour reconstruire la phase. Par exemple j'ai développé un algorithme utilisant comme variables intermédiaires de la phase (slacks variables), les amplitudes complexes de l'onde EM dans les deux plans de Fourier conjugués de la pupille et de l'image, et les mettant à jour itérativement, l'une après l'autre. J'ai en fait principalement ajouté une contrainte de lissage

dans l'espace de la pupille pour pouvoir déplier la phase, par rapport aux formulations de type POCS proposées par Fienup [1982]. Bien que cette approche avec des slacks variables m'ait dans certains cas permis de reconstruire la phase dans un temps de calcul moindre par rapport à la formulation du maximum de vraisemblance a posteriori, l'erreur de reconstruction de la phase n'a jamais été aussi bonne qu'avec le maximum a posteriori. J'ai donc tenté d'utiliser les slacks variables en tant que stratégie d'optimisation globale servant à sélectionner le bon minimum local, avant raffinement par maximum de vraisemblance. Mais je n'ai en pratique pas su développer un algorithme de slacks variables pouvant s'adapter à des conditions de turbulence et de bruit très diverses avec autant de succès que l'algorithme détaillé dans ce chapitre.

Une alternative peut-être moins compliquée pour initialiser l'algorithme d'optimisation dans le bon minimum local, mais que je n'ai cependant pas eu le temps de mettre en oeuvre également, me semble de rééchantillonner l'image parallèlement à une plus forte restriction du nombre initial de paramètres de phase, et cela afin de ne pas ajuster de haute fréquence spatiale de l'image sur les mauvais modes de phase.

Enfin, un travail d'optimisation des différents ingrédients de l'algorithme d'optimisation globale détaillé dans ce chapitre, en fonction des conditions de turbulence et de bruit, et en utilisant la même démarche de caractérisation que celle que j'ai développée, me semble devoir être fait pour pouvoir au besoin améliorer encore les performances de l'algorithme.

Chapitre 4

Diversité de phase polychromatique

La reconstruction de la phase sur la pupille d'un télescope à partir d'une seule image tavelée nécessite une stratégie d'optimisation globale efficace en plus d'un algorithme d'optimisation locale, comme on l'a vu au chapitre précédent. Les deux stratégies que j'ai développées ne sont toutefois plus assez performantes à partir de $D/r_0 \simeq 13$, à cause des ambiguïtés liées à la mauvaise estimation du tip-tilt par le barycentre notamment.

Au-delà de son emploi premier pour lever l'ambiguïté du signe de la partie paire de la phase, la diversité focale est une technique éprouvée pour simplifier la reconstruction de la phase [Gonsalves, 1982; Idier et al., 2005]. Elle n'a toutefois été envisagée qu'assez récemment [Mugnier et al., 2006], pour estimer des phases turbulentes avec un grand nombre de degrés de liberté, et à ma connaissance dans le cadre particulier des corrections de phase a posteriori seulement, qui ne nécessitent pas de dépliage de la phase.

Dans le cadre du projet ELP-OA, je m'intéresse ici à une diversité de phase alternative appelée diversité de phase *polychromatique*, dont le concept a été présenté dans [Vaillant et al., 2000; Vaillant, 2002]. Cette diversité consiste à considérer un ensemble d'images monochromatiques $m_\lambda(\mathbf{x})$ pour reconstruire le même chemin optique $\delta(\mathbf{u})$, auquel les phases aux différentes longueurs d'onde $\varphi_\lambda(\mathbf{u})$ sont reliées par un simple scalaire chromatique.

Après avoir développé un modèle polychromatique de reconstruction de phase, je montre dans ce chapitre que la diversité polychromatique de phase permet de gérer des forces de turbulence (D/r_0) beaucoup plus importantes que dans le cas monochromatique. Pour automatiser l'algorithme polychromatique par rapport à des conditions de turbulence et de bruit très variées, je présente ensuite une stratégie d'ajustement automatique de la pondération de chaque critère expérimental monochromatique f_{exp,λ_i} dans la fonction de pénalisation polychromatique $f_{\text{post}}(\mathbf{p})$, dont on cherche le minimum global pour estimer le chemin optique.

4.1 Modélisation polychromatique

4.1.1 Modèle des images

Le modèle direct continu des images a été dérivé à la section 1.1.2.2. La différence ici par rapport à la modélisation monochromatique est que les phases $\varphi_{\lambda_i}(\mathbf{r})$ sont ré-exprimées en fonction d'une unique phase de référence $\varphi_{\lambda_1}(\mathbf{r})$, à l'aide de la relation entre deux phases dérivée à la section 1.4.4.2 :

$$\varphi_{\lambda_i}(\mathbf{r}) = \frac{\lambda_1 (n_{\lambda_i} - 1)}{\lambda_i (n_{\lambda_1} - 1)} \times \varphi_{\lambda_1}(\mathbf{r}) \quad (4.1)$$

$$= \Upsilon_{\lambda_{i1}} \times \varphi_{\lambda_1}(\mathbf{r}). \quad (4.2)$$

En effet, le chemin optique qui résulte de la traversée de l'atmosphère est à peu près le même pour les différentes longueurs d'onde, ce qui permet de ne conserver qu'un seul jeu de paramètres de phase pour modéliser les images $m_{\lambda_i}(\mathbf{x})$ aux différentes longueurs d'onde.

Comme on l'a vu lors de la discrétisation du modèle direct des images dans le cas monochromatique (cf. section 3.2.1), le calcul de l'image discrète \mathbf{m}_{λ_i} , nécessite la transformée de Fourier discrète de la phase

$$\varphi_{\lambda_i,k} \triangleq \varphi_{\lambda_i}(\lambda_i \mathbf{u}_k) \quad (4.3)$$

La réexpression de toutes les phases discrètes en fonction de $\varphi_{\lambda_1,k} \triangleq \varphi_{\lambda_1}(\lambda_1 \mathbf{u}_k)$, amène donc en principe à devoir rééchantillonner $\varphi_{\lambda_1,k}$ par le facteur de grandissement λ_i/λ_1 pour pouvoir calculer chaque image \mathbf{m}_{λ_i} . Alternativement, on peut supposer que ce rééchantillonnage est effectué optiquement, par l'introduction des grandsissements appropriés sur les canaux optiques formant les images $\mathbf{m}_{\lambda_{i \neq 1}}$. Cela permet en effet d'une part que les tavelures — i.e. l'information haute fréquence des images — soient échantillonnées avec la même finesse pour les différentes longueurs d'onde sur le CCD, lorsque le même CCD détecte les images spatialement éclatées aux différentes longueurs d'onde notamment. Cela permet également de se passer de rééchantillonneur dans la modélisation discrète, et d'éviter ainsi des erreurs d'interpolation lors de la reconstruction. Par la suite je considère donc que le rééchantillonnage nécessaire est effectué optiquement. Toutes les images discrètes \mathbf{m}_{λ_i} sont alors modélisées à l'identique, aux facteurs scalaires près $\Upsilon_{\lambda_{i1}}$ pour les phases sur la pupille, et α_{λ_i} pour les intensités des images.

4.1.2 Critère d'adéquation a posteriori polychromatique

En faisant l'hypothèse que les fluctuations des modèles monochromatiques sont liées au bruit de photon et au bruit de détecteur, on considère que les modèles monochromatiques sont statistiquement indépendants entre eux. Ainsi le critère expérimental d'optimisation polychromatique — qui est le logarithme négatif de la densité de

probabilité expérimentale polychromatique — devrait en principe s'écrire comme la somme des critères monochromatiques :

$$f_{\text{exp}}(\mathbf{p}) = \sum_i f_{\text{exp},\lambda_i}(\mathbf{p}), \quad (4.4)$$

pourvu que la pondération de chaque critère par la matrice de covariance du bruit soit correcte. Pour favoriser l'un des critères monochromatiques lors de l'optimisation, en fonction du nombre de ses ambiguïtés, j'exprime en fait le critère expérimental polychromatique sous la forme

$$f_{\text{exp}}(\mathbf{p}) = \sum_i \mu_i f_{\text{exp},\lambda_i}(\mathbf{p}), \quad (4.5)$$

où les μ_i sont des hyperparamètres de pondération qui peuvent être réglés à chaque pas de l'optimisation.

En conservant la même régularisation et paramétrisation que dans le cas monochromatique, le critère d'adéquation a posteriori polychromatique s'écrit au final

$$f_{\text{post}}(\boldsymbol{\phi}, \boldsymbol{\theta}) = \sum_{i,j} \mu_i \times \left(\frac{d_{\lambda_i,j} - \alpha_{\lambda_i}^\dagger m_{\lambda_i,j}(\boldsymbol{\phi}, \boldsymbol{\theta})}{\sigma_{\lambda_i,j}} \right)^2 + \boldsymbol{\varphi}'^\top \cdot \mathbf{C}_{\boldsymbol{\varphi}'}^{-1} \cdot \boldsymbol{\varphi}'$$

4.1.3 Algorithme d'optimisation

L'algorithme d'optimisation polychromatique que j'utilise pour la reconstruction de phase polychromatique est le même que celui que j'ai développé et détaillé dans le cas monochromatique. Les directions du sous-espace d'optimisation que je considère, sont celles issues de chaque critère expérimental monochromatique indépendamment, comme dans la section 3.5.3. La séparation des directions entre les différents critères permet de laisser la flexibilité de réajustement des poids des critères, lorsque les hyperparamètres μ_i sont ajustés à chaque itération, et que ceux disponibles pour le calcul des directions sont ceux déterminés à l'itération précédente.

Par ailleurs, j'utilise la même reparamétrisation de la phase à l'aide des noeuds de spline d'un interpolateur bicubique, en considérant cette fois-ci un nombre de noeuds égal à $\sim (D/r_0(\lambda_{\min}))^2$ pour pouvoir s'attacher aux structures fines de toutes les images.

4.2 Bras de levier polychromatique

La difficulté de la reconstruction monochromatique de la phase à partir d'une seule image tavelée, en terme de nombre de minima locaux, augmente fortement avec le D/r_0 et avec les augmentations parallèles des amplitudes spatiales de la phase comme on l'a vu au chapitre précédent. Les critères monochromatiques $f_{\text{exp},\lambda_i}(\mathbf{p})$ du critère

polychromatique, ne comportent donc pas le même nombre de minima locaux aux différentes longueurs d'onde, car le D/r_0 est proportionnel à $\lambda^{-1.2}$. La modélisation est en fait tellement non-linéaire qu'on peut aussi considérer que les minima locaux sont totalement décorrélés et sont situés à des endroits différents entre les différents critères.

C'est pourquoi on peut penser s'appuyer comme un bras de levier sur les critères les plus faciles, c'est-à-dire ceux aux longueurs d'onde les plus longues, pour gommer un nombre trop important de minima locaux dans les premiers temps de l'optimisation. Par exemple, il semble plus judicieux de renforcer l'hyperparamètre qui pondère un critère expérimental dans le proche infra-rouge, où D/r_0 est relativement faible, par rapport au critère dans le visible, pour lequel le D/r_0 est plus élevé.

J'ai vérifié l'efficacité du bras de levier polychromatique par des simulations avec un couple de deux longueurs d'onde ($\lambda_2 < \lambda_1$) et sur un ensemble de phases temporellement décorrélées. J'ai considéré pour cela un $D/r_0(\lambda_1) = 8$ constant et j'ai progressivement diminué la seconde longueur d'onde pour augmenter le $D/r_0(\lambda_2)$ de 18 jusqu'à 69, ce qui correspond finalement à un couple de longueurs d'onde de 0.3 et 2.6 μm environ. Les nombres de photons considérés sont de $N_{\text{ph1}} = 10^5$ et $N_{\text{ph2}} = [5 \times 10^4, 1.36 \times 10^5, 3 \times 10^5, 5.1 \times 10^5, 7.27 \times 10^5]$ photons pour $D/r_0(\lambda_2) = [18, 30, 45, 58, 69]$, ce qui correspond à un rapport signal à bruit par tavelures de 35 pour λ_1 et 11 pour λ_2 .

Dans un premier temps, j'ai ajusté les poids des critères monochromatiques de façon un peu brutale, en vérifiant avant tout l'idée de bras de levier. J'ai ainsi choisi d'assurer la convergence dans le bon minimum global de $f_{\text{exp}, \lambda_1}(\mathbf{p})$, avec $\mu_2 = 0$ au début, et en utilisant la stratégie d'optimisation globale avec les départs aléatoires, qui a été décrite au chapitre précédent. Puis de rajouter la seconde longueur d'onde à l'issue de cette première étape, et d'augmenter progressivement μ_2 toutes les $\simeq 20$ itérations, jusqu'à la convergence finale de l'optimisation de $f_{\text{post}}(\mathbf{p})$.

Comme on le voit sur les courbes de la figure 4.2, j'ai pu reconstruire des phases avec des D/r_0 jusqu'à 69 en utilisant la stratégie que je viens de décrire. De par la remise à l'échelle $\Upsilon_{\lambda_{21}}$, l'erreur de reconstruction de la phase ϕ_{λ_2} est multipliée par $\Upsilon_{\lambda_{21}}$ par rapport à celle de reconstruction de ϕ_{λ_1} . Ainsi, il doit arriver un moment où la convergence de la première phase ne permet plus de contraindre suffisamment bien la deuxième longueur d'onde par rapport à son nombre élevé de minima locaux. En pratique la limite de $D/r_0(\lambda_2)$ à laquelle je me suis arrêtée pour les simulations, a plutôt été causée par l'insuffisance de la puissance de calcul et de stockage permettant de gérer un nombre plus grand de degrés de liberté, que par un début de défaillance du bras de levier.

La comparaison des différentes courbes révèle que la précision de reconstruction a été influencée par l'hétérogénéité d'ajustement des hyperparamètres avec le D/r_0 . Cela m'a conduit à mettre en place une stratégie plus élaborée pour ajuster les hyperparamètres μ_i , comme cela est détaillé dans la section suivante.

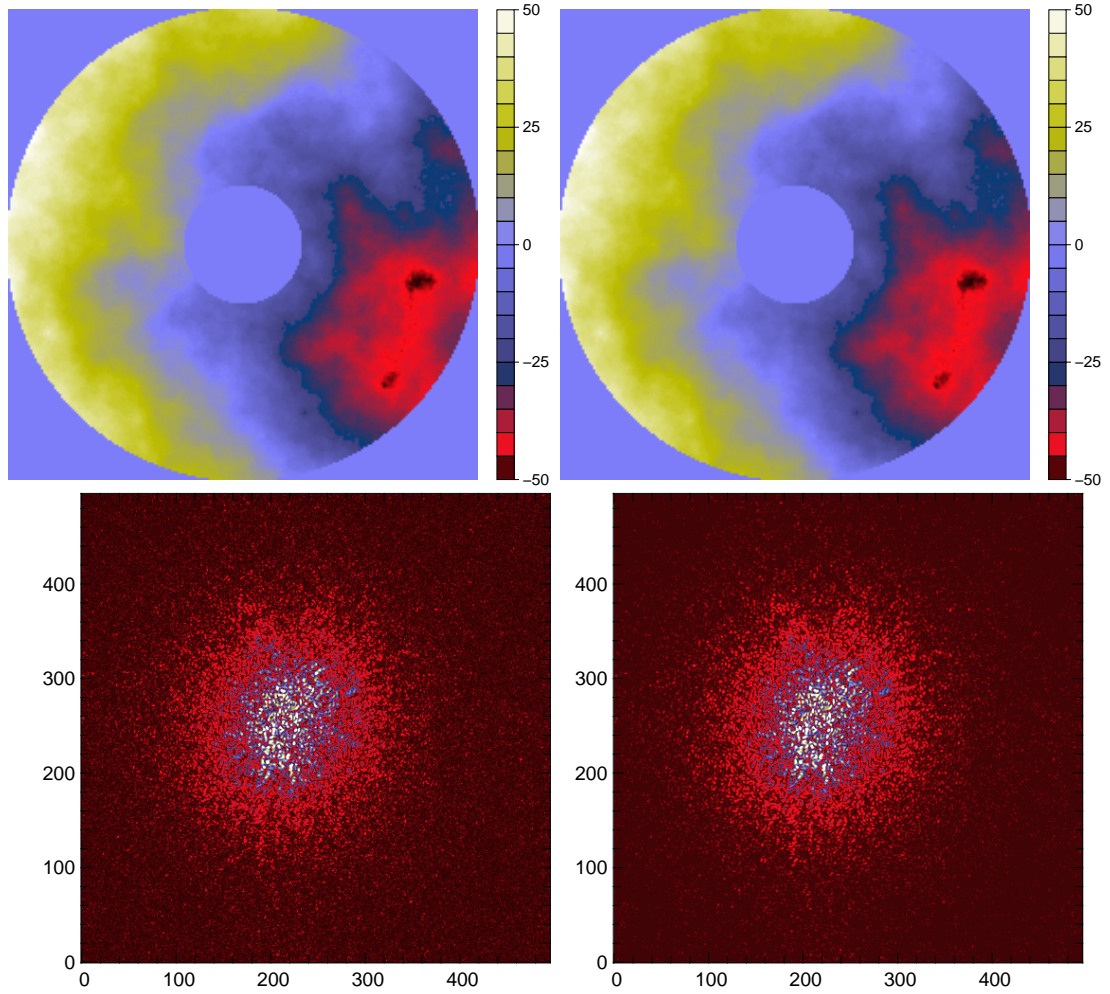


FIG. 4.1 — Exemple de phases simulée (en haut à gauche) et reconstruite (à droite) sans aucune approximation grossière initiale de la phase et pour $D/r_0 = 58$. Les images tavelées correspondantes sont représentées en bas. La palette des amplitudes de phase est en radian.

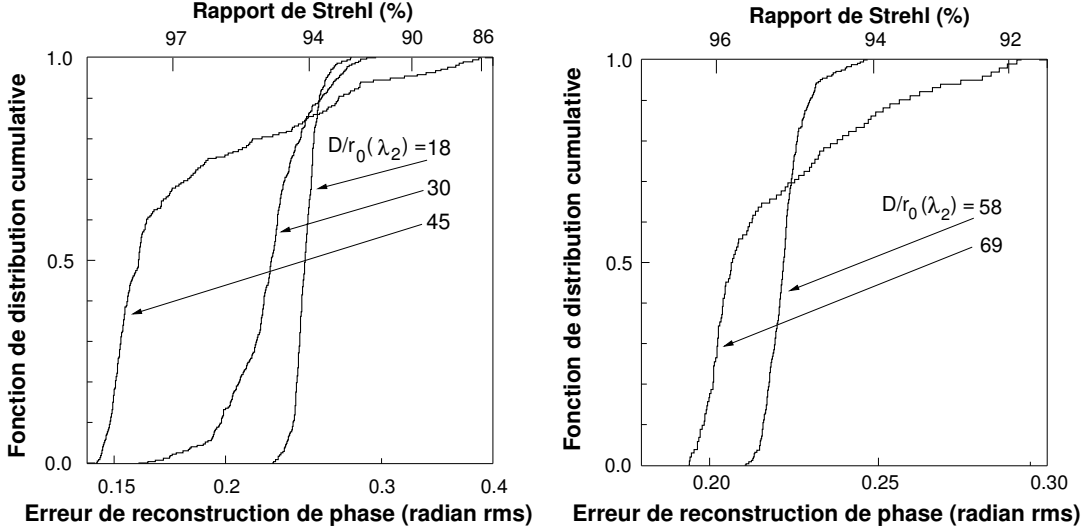


FIG. 4.2 — Erreur de reconstruction de la phase φ_{λ_2} par diversité polychromatique de phase, avec $D/r_0(\lambda_1)=8$

4.3 Ajustement des poids relatifs des critères monochromatiques

Puisque certains critères expérimentaux de la fonction a posteriori sont plus lisses que d'autres — au même titre que la régularisation —, il m'a semblé assez naturel d'étendre la stratégie d'ajustement du poids de la régularisation décrite à la section 3.5.2 à tous les hyperparamètres μ_i . En principe cela suppose d'optimiser simultanément tous les hyperparamètres, par rapport à une certaine proportionnalité à obtenir entre les variations avec et sans les hyperparamètres.

Pour simplifier cet ajustement simultané de plusieurs hyperparamètres, qui de plus, doit avoir lieu à chaque itération, j'ai séparé de façon cascadée l'ajustement de chaque poids. Ainsi pour le cas de reconstruction avec deux longueurs d'onde par exemple, μ_1 est ajusté en premier pour que la régularisation $f_{\text{prior}}(\mathbf{p})$ contraigne suffisamment les variations de $f_{\text{exp},\lambda_1}(\mathbf{p})$:

$$\delta f_{\text{exp},\lambda_1}^{\text{quad}}(\delta \mathbf{p}_{\text{post}}(\mu_1, \mu_2=0)) \simeq \epsilon_1 \times \delta f_{\text{exp},\lambda_1}^{\text{quad}}(\delta \mathbf{p}_{\text{exp}}(\mu_1, \mu_2=0)) ; \quad (4.6)$$

avec $0 < \epsilon_1 < 1$, et

$$\begin{aligned} \delta \mathbf{p}_{\text{exp}}(\mu_1, \mu_2) &= - \left[\mu_1 \times \mathbf{H}_{\text{exp},\lambda_1} + \mu_2 \times \mathbf{H}_{\text{exp},\lambda_2} + \ell \times \mathbf{D} \right]^{-1} \cdot (\mu_1 \times \mathbf{g}_{\text{exp},\lambda_1} + \mu_2 \times \mathbf{g}_{\text{exp},\lambda_2}) \\ \delta \mathbf{p}_{\text{post}}(\mu_1, \mu_2) &= - \left[\mathbf{H}_{\text{prior}} + \mu_1 \times \mathbf{H}_{\text{exp},\lambda_1} + \mu_2 \times \mathbf{H}_{\text{exp},\lambda_2} + \ell \times \mathbf{D} \right]^{-1} \cdot (\mathbf{g}_{\text{prior}} + \mu_1 \times \mathbf{g}_{\text{exp},\lambda_1} + \mu_2 \times \mathbf{g}_{\text{exp},\lambda_2}) . \end{aligned}$$

Puis μ_2 est ensuite ajusté pour que les variations de $f_{\text{exp},\lambda_2}(\mathbf{p})$ soient suffisamment guidées par celles de $f_{\text{prior}}(\mathbf{p}) + \mu_1 \times f_{\text{exp},\lambda_1}(\mathbf{p})$:

$$\delta f_{\text{exp},\lambda_2}^{\text{quad}}(\delta \mathbf{p}_{\text{post}}(\mu_1, \mu_2)) \simeq \epsilon_2 \times \delta f_{\text{exp},\lambda_2}^{\text{quad}}(\delta \mathbf{p}_{\text{exp}}(\mu_1=0, \mu_2)), \quad (4.7)$$

avec $0 < \epsilon_2 < 1$.

Cette stratégie de prise en compte progressive des critères les plus multi-modaux s'est révélée tout aussi efficace que dans le cas monochromatique, pour gérer automatiquement diverses conditions de turbulence et de bruit. Comme on le voit sur la figure 4.3, on obtient en effet un comportement plus uniforme de l'erreur de reconstruction avec la variation de l'écart du couple de longueurs d'onde (à RSB par tavelure constant entre les 2 courbes). De plus il faut noter que les ϵ_i peuvent par ailleurs être optimisés de façon à combiner l'information portée par les images aux différentes longueurs d'onde, au mieux pour l'optimisation, c'est à dire pour avoir le moins possible de minima locaux.

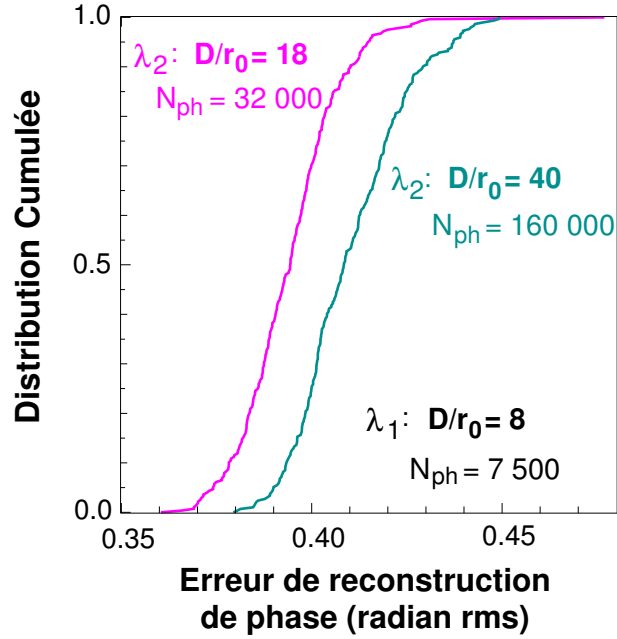


FIG. 4.3 — Erreur de reconstruction de la phase φ_{λ_2} par diversité polychromatique de phase, avec un meilleur ajustement des poids relatifs μ_i

4.4 Conclusion

J'ai présenté dans ce chapitre un prolongement aux développements de reconstruction de phase monochromatique du chapitre précédent, en exploitant le bénéfice de

la diversité polychromatique de phase pour reconstruire des phases avec des rapports de D/r_0 beaucoup plus élevés que dans le cas monochromatique. Ainsi je parviens à gérer de l'ordre de 5000 degrés de liberté ($D/r_0=69$), contre 120 dans le cas monochromatique. J'ai par ailleurs développé une méthode d'ajustement des poids des critères monochromatiques, qui me permet de combiner de façon plus judicieuse l'information aux différentes longueurs d'onde pour éviter de trop nombreux minima locaux. Le gain obtenu en terme de degrés de liberté pour un système d'Optique Adaptative est supérieur à celui obtenu jusqu'à présent dans ce contexte par la technique de diversité focale de phase [Mugnier et al., 2006], et cela grâce au principe intrinsèque de bras de levier polychromatique et également grâce au bénéfice de la combinaison des longueurs d'onde (accord de chemin optique) pour déplier la phase. L'emploi de plusieurs longueurs d'onde pour estimer des périodes de phase est en effet bien connu des opticiens, depuis que Michelson and Benoit [1895] ont effectué les premières calibrations du mètre à l'aide de la méthode des excédents fractionnaires et de plusieurs longueurs d'onde. Il se révèle également tout à fait bien adapté, en combinaison avec le bon a priori sur la covariance spatiale de la phase, pour déterminer sans ambiguïté la périodicité des échantillons de la phase à partir des images, dans le cas de très forts D/r_0 (amplitudes de phase de quelques centaines de radian sur la pupille) et cela bien que l'information de la période de chaque échantillon de phase ne soit pas contenue dans les images monochromatiques prises séparément.

Dans le cadre du projet ELP-OA, la reconstruction de la phase pour de tels D/r_0 pourrait permettre un gain très important sur la précision de l'estimation de la position de l'étoile laser, typiquement de l'ordre du rapport de taille entre la tache image de turbulence et la tache image de diffraction, ce qui est démontré et discuté plus en détail au chapitre suivant.

Une perspective intéressante pour les algorithmes de reconstruction de phase monochromatique aussi bien que polychromatique me semblerait de trafiquer la métrique de la fonction d'adéquation expérimentale pour diminuer le nombre de minima locaux. En effet on a vu que l'ajout des deux critères monochromatiques dans l'adéquation polychromatique selon leur poids statistique conduit à un plus grand nombre de minima locaux que dans le cas monochromatique pour la longueur d'onde la plus complexe à traiter, même si l'approche devient très robuste localement. De même dans le cas monochromatique où la formulation gaussienne est robuste au voisinage du minimum global, rien ne dit que la métrique quadratique soit la plus judicieuse vis à vis du nombre de minima locaux.

Chapitre 5

Centrage

Dans ce chapitre je regroupe les différents travaux de centrage que j'ai effectués. Ils concernent l'amélioration de l'estimation de la position d'un objet observé à travers la turbulence d'une part, et l'amélioration de la résolution spatiale d'un détecteur à comptage de photons d'autre part. Je montre tout d'abord que la position d'une source ponctuelle peut être estimée plus précisément que par un simple barycentre des images tavelées, lorsque toutes les aberrations de la phase sur la pupille, dues à la traversée du milieu turbulent, sont conjointement estimées. Je présente dans la deuxième section une nouvelle méthode d'estimation de la pente chromatique différentielle de la surface d'onde par interspectre bichromatique. Dans le cadre du projet d'Etoile Laser Polychromatique, je compare les performances de cette méthode que l'objet soit résolu ou non par le télescope. Enfin je présente une technique pour améliorer la résolution spatiale d'un détecteur à comptage de photons, et je montre ses bénéfices concrets sur de vraies données.

5.1 Estimation de la position d'une source ponctuelle à partir d'images tavelées

5.1.1 Différence entre barycentre et pente de la surface d'onde

Dans le cas où l'amplitude de l'onde ne fluctue pas sur la pupille, on peut considérer la pente du chemin optique sur la pupille au lieu du barycentre des images, pour exprimer le déplacement de l'objet observé. En effet en utilisant la propriété de translation de la transformée de Fourier \mathcal{F} :

$$\mathcal{F} \left[G(u) \times e^{2\pi j t u} \right] (x) = \mathcal{F} [G(u)] (x - t), \quad (5.1)$$

et en l'associant à l'expression des images au foyer du télescope dérivée à la section 1.1.2.2 :

$$i_\lambda(\mathbf{x}) = \frac{1}{(\lambda f)^2} \times \left| \mathcal{F} \left[\left(P_\lambda(\mathbf{u}) e^{2\pi j \times \delta_\lambda(\mathbf{u})} \right) \right] \right|^2, \quad (5.2)$$

on voit que les coefficients de pente $[\theta_1, \theta_2]$ expriment un déplacement de l'objet dans l'image, tout comme le barycentre.

D'autre part, le barycentre des images se réexprime comme le gradient moyen du chemin optique (cf. section 1.4.4.2). Au fur et à mesure que les aberrations autres que la pente augmentent avec le D/r_0 dans $\delta_\lambda(\mathbf{u})$, le gradient moyen et la pente du chemin optique deviennent assez différents. Le barycentre et les coefficients de pente n'expriment donc pas le même déplacement de l'objet. En décomposant la phase dans une base modale de Zernike, on remarque que les modes qui ajoutent du déplacement dans le gradient moyen par rapport au déplacement de la pente sont les modes centro-dissymétriques ou modes impairs. Ce sont en particulier les premiers modes impairs, puisqu'ils ont une variance plus élevée, qui causent cette différence dans le gradient moyen (la coma notamment).

En utilisant une turbulence de type Kolmogorov et une pupille circulaire, on montre [Sasiela, 1994] que l'écart-type de la distribution temporelle de la différence selon un axe entre la pente de la surface d'onde et le barycentre vaut :

$$\sigma_{\theta_k - b_k}^{\text{temp}} = 0.055 \times \frac{\lambda}{D} \times \left(\frac{D}{r_0} \right)^{5/6}. \quad (5.3)$$

De plus les écart-types des distributions temporelles du tip-tilt θ et du barycentre \mathbf{b} valent sous ces mêmes hypothèses

$$\begin{aligned} \sigma_{\theta_k}^{\text{temp}} &\simeq 0.427 \times \frac{\lambda}{D} \times \left(\frac{D}{r_0} \right)^{5/6} \\ \sigma_{b_k}^{\text{temp}} &\simeq 0.412 \times \frac{\lambda}{D} \times \left(\frac{D}{r_0} \right)^{5/6}. \end{aligned} \quad (5.4)$$

L'augmentation de la différence entre le tip-tilt et le barycentre en fonction du D/r_0 est représentée sur la figure 5.1. On voit que le tip-tilt ne peut pas être bien estimé à partir du barycentre des images et vice versa. En pratique l'instrumentation introduit également un décalage supplémentaire entre le barycentre et le tip-tilt, mais ce dernier est calibrable et peut donc être compensé.

5.1.2 Erreurs et biais affectant le centrage

J'exprime ici les erreurs et biais affectant le centrage des images tavelées pour le barycentre uniquement, puisqu'il n'existe pas de relation analytique linéaire de la pente de la surface d'onde en fonction de l'énergie dans l'image, et puisque je m'intéresse par la suite à des techniques de centrage à partir de l'énergie de l'image. Dans le cas

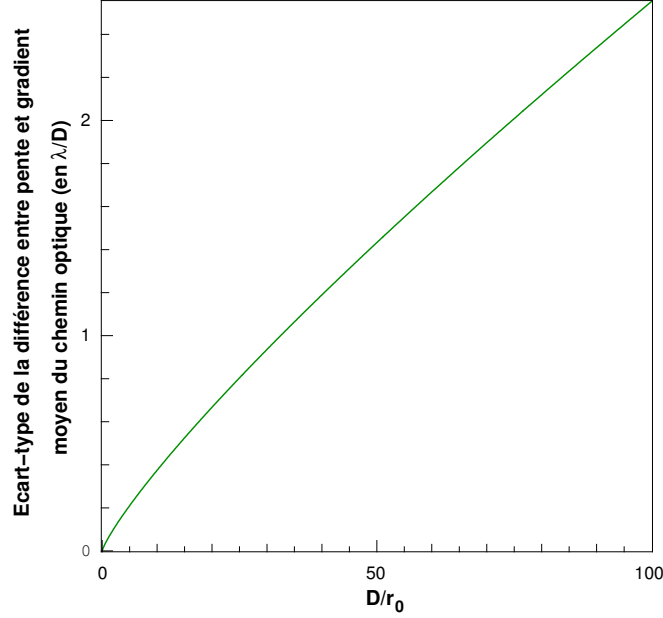


FIG. 5.1 – Ecart-type de la différence entre la pente du chemin optique et le barycentre de l'image

où le bruit de photons domine, je suppose que l'erreur d'un estimateur de maximum de vraisemblance du barycentre est du même ordre de grandeur que celle d'un estimateur de maximum de vraisemblance de la pente du chemin optique (tip-tilt).

5.1.2.1 Bruit de discrétisation spatiale

Le déplacement continu de la position de l'objet est estimé par une version discrète du barycentre, car les images formées par le détecteur sont spatialement discrètes. En notant Δx la taille des pixels de l'image, la version discrète du barycentre s'écrit

$$b_k = \frac{\sum_{l,m} (x_{k,l,m} \Delta x) \times i_{l,m}}{\sum_{l,m} i_{l,m}} \quad k = 1, 2. \quad (5.5)$$

On montre [Vaillant, 2002] que la discrétisation des images n'introduit pas de biais dans l'estimation du barycentre, mais introduit une erreur

$$\sigma_{b_k}^{\text{disc}} = \frac{\Delta x}{\sqrt{12 \times N_{\text{ph}}}}, \quad (5.6)$$

où N_{ph} est le nombre de photons dans l'image. Cette erreur est le plus souvent négligeable devant les erreurs dues au bruit de photons et au bruit du détecteur. En pratique la discrétisation spatiale peut entraîner une erreur voire également un biais, lorsque la transmission des pixels est fortement hétérogène et qu'elle n'est pas calibrée et compensée.

5.1.2.2 Bruit du détecteur

En dehors des troncatures spatio-temporelles, le détecteur modifie la distribution spatiale d'énergie du signal photonique, à cause principalement du courant thermique du matériau optoélectronique, du transfert des charges électroniques (la lecture) et de la digitalisation du signal. Après débiaisage par soustraction d'un champ obscur moyenné sur un nombre significatif de trames, on considère en général que le bruit du détecteur est spatialement homogène et sans biais. Chaque pixel du détecteur est ainsi affecté par un bruit gaussien centré d'écart type temporel σ_{CCD} . En général le bruit de lecture domine le bruit du CCD, de telle sorte que σ_{CCD} est proportionnel à la racine carrée de la fréquence d'acquisition. L'estimateur du barycentre en présence du bruit de détecteur $n_{l,m}$ sur chaque pixel s'écrit

$$\tilde{b}_k = \frac{\sum_{l,m} (x_{k,l,m} \Delta x) \times (i_{l,m} + b_{l,m})}{\sum_{l,m} (i_{l,m} + n_{l,m})}. \quad (5.7)$$

Lorsque le flux est relativement plus important que le bruit de lecture, on considère que la somme du dénominateur de l'équation (5.7) n'est pas affectée par le bruit de photons. On voit alors directement de la moyenne de la soustraction de cette équation avec l'équation (5.5) que l'estimateur du barycentre n'est pas biaisé à la condition que la fenêtre de centrage soit centrée sur le barycentre. Pour réaliser cette condition, on peut par exemple utiliser un algorithme de barycentre itératif dont la fenêtre de centrage est recentrée à chaque itération sur le barycentre calculé à l'itération précédente.

En développant l'écart quadratique moyen entre les équations (5.5) et (5.7), on voit également que le bruit homogène du CCD sur chaque pixel se traduit dans l'estimation du barycentre par l'erreur

$$\sigma_{b_k}^{\text{det}} = \frac{\sigma_{\text{CCD}}}{N_{\text{ph}}} \times \sqrt{\sum_{l,m} (k_{l,m} \Delta x)^2}. \quad (5.8)$$

L'erreur de centrage liée au bruit du détecteur augmente donc avec la taille et le nombre des pixels. C'est pourquoi on choisit souvent un recouvrement du signal sur 2×2 pixels du détecteur, à la limite de diffraction, pour estimer le barycentre lorsque le bruit de lecture domine. Dans ce cas-là toute troncature du signal est dramatique pour le barycentre, comme cela est illustré sur la figure 5.19.

5.1.2.3 Bruit de photons

Le bruit de photons est la principale source d'erreur pour le centrage dans le cas d'acquisitions à faibles flux, et/ou en faisant l'anticipation optimiste de l'amélioration technologique des détecteurs et de la diminution conséquente de leur bruit. Ce bruit est dû à la nature quantique de l'interaction de la lumière avec la matière. En effet,

comme le détecteur intègre de l'énergie sur un intervalle de temps fini, sous la forme de paquets d'énergie, la moyenne instantanée de l'énergie acquise par le détecteur sur une trame est fluctuante pour un flux de lumière constant.

Soit $h(\mathbf{x})$ la distribution d'énergie dans le plan focal du détecteur, que l'on considère normalisée et recentrée au milieu du détecteur. En supposant que $h(\mathbf{x})$ est une gaussienne, on montre [Lane et al., 2005] que l'erreur d'estimation de la position de la distribution, liée au bruit de photons, vaut (en unité angulaire)

$$\sigma_{b_k}^{\text{ph}} = \frac{0.657 \times \lambda/D}{\sqrt{N_{\text{ph}}}}. \quad (5.9)$$

Lorsque le calcul est mené avec une distribution d'Airy plutôt qu'avec une distribution gaussienne, on montre [van Dam and Lane, 2002b] que l'erreur d'estimation du barycentre liée au bruit de photons est infinie. Ce résultat surprenant est dû au fait que la distribution d'Airy décroît beaucoup plus lentement spatialement qu'une gaussienne. Il n'existe pas de relation analytique pour les distributions d'énergie des images tavelées, puisque celles-ci sont intrinsèquement fluctuantes, de sorte qu'on ne peut pas calculer l'erreur sur le barycentre liée au bruit de photons. Toutefois comme la distribution moyenne des images courtes poses recentrées décroît plutôt comme la distribution gaussienne que comme la distribution d'Airy pour les forts $D/r_0 \geq 10$, on s'attend à ce que l'erreur de barycentre liée au bruit de photons pour les images tavelées à fort D/r_0 soit finie, et analogue à celle de la gaussienne, en remplaçant la taille de la gaussienne dans l'équation (5.9) par celle de la distribution d'intensité moyenne des images tavelées (courtes poses) recentrées. Dans un régime de D/r_0 intermédiaire de 2 à 10, on s'attend par contre à ce que l'erreur du barycentre soit plus importante, parce que la décroissance spatiale de la distribution moyenne d'intensité est alors plus lente que celle d'une gaussienne.

5.1.2.4 Troncature par le détecteur

La taille finie du détecteur introduit un biais et une erreur de centrage vers le centre du détecteur, parce que la distribution d'énergie de l'image est tronquée de façon dissymétrique. Le biais et l'erreur augmentent d'autant plus que le détecteur n'est pas placé au milieu de la distribution d'intensité.

La troncature comporte toutefois certains avantages et peut dans certains cas être intentionnellement renforcée, en choisissant une fenêtre de calcul du barycentre inférieure à la taille du détecteur. En effet la troncature introduit tout d'abord un a priori qui permet de rendre finie l'erreur d'estimation du barycentre liée au bruit de photons pour la distribution d'Airy et aussi de la limiter pour les images tavelées à faible D/r_0 . La troncature permet également d'améliorer l'estimation du barycentre en présence d'un bruit additif comme le bruit de lecture du CCD par exemple, car les pixels des bords de la distribution d'énergie, où le niveau du bruit domine le signal, ne sont ainsi pas pris en compte.

5.1.3 Borne de Cramér-Rao

La borne de Cramér-Rao est une limite inférieure pour l'erreur d'estimation d'un paramètre p , lorsque p est estimé par maximum de vraisemblance à partir de mesures \mathbf{x} distribuées selon la densité de probabilité $h(\mathbf{x}|p)$:

$$p^+ = \arg \max_p h(\mathbf{x}|p). \quad (5.10)$$

On montre [Cramér, 1946] que cette borne est liée à la courbure de la densité de probabilité $h(\mathbf{x}|p)$:

$$\begin{aligned} \sigma_p^{+2} &= - \left(E \left[\frac{\partial^2 \ln(h(\mathbf{x}|p))}{\partial p^2} \right] \right)_{|p=p^+}^{-1} \\ &= \frac{1}{I(p^+)}, \end{aligned} \quad (5.11)$$

où $I(p)$ est appelée l'information de Fisher. Dans le cas multidimensionnel, la borne de Cramér-Rao s'exprime par

$$\mathbf{C}_p^+ - \mathbf{I}_p^{-1} \quad \text{positive définie}, \quad (5.12)$$

où $\mathbf{C}_p = \langle \mathbf{p}, \mathbf{p}^\top \rangle$ est la matrice de covariance de \mathbf{p} , et \mathbf{I}_p est la matrice d'information de Fisher.

Lorsque le bruit de photons est le bruit dominant dans les images, on peut considérer que chaque mesure de photon possède sa propre fluctuation indépendante [van Dam and Lane, 2000]. L'estimateur de maximum de vraisemblance du déplacement \mathbf{b} ou θ , que l'on obtient à partir des positions \mathbf{x}_i des photons mesurées dans l'image, s'écrit donc

$$\mathbf{b}^+ = \arg \max_{\mathbf{b}} \prod_i h(\mathbf{x}_i|\mathbf{b}). \quad (5.13)$$

En supposant que \mathbf{I}_b ne dépend pas de \mathbf{x}_i , et puisque chaque mesure de position de photon \mathbf{x}_i est indépendante et combinée aux autres pour estimer le même paramètre \mathbf{b} , la borne de Cramér-Rao s'écrit

$$\mathbf{C}_b^+ - \frac{1}{N_{\text{ph}}} \times \mathbf{I}_b^{-1} \quad \text{positive définie}. \quad (5.14)$$

Pour une distribution d'intensité gaussienne, on a

$$h(\mathbf{x}|\mathbf{b}) = \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp \left[-(\mathbf{x} - \mathbf{b})^2 / (2\sigma_g^2) \right], \quad (5.15)$$

En calculant la matrice de Fisher, et en observant que celle-ci est diagonale, il vient

$$\sigma_{b_k}^+ = \frac{\sigma_g}{\sqrt{N_{\text{ph}}}}, \quad (5.16)$$

pour chaque composante vectorielle $k = 1, 2$ du déplacement.

Pour une distribution d'Airy associée à une pupille de diamètre D , on montre [van Dam and Lane, 2000] que la borne de Cramér-Rao selon chaque axe vaut (en unité angulaire) :

$$\sigma_{b_k}^+ = \frac{0.318 \times \lambda/D}{\sqrt{N_{\text{ph}}}} \quad (5.17)$$

Cette borne est inférieure à l'erreur d'estimation du barycentre d'une gaussienne, exprimée à l'équation (5.9) lorsque le bruit de photons domine. De plus, le barycentre est un estimateur de maximum de vraisemblance du déplacement pour une gaussienne, mais pas pour une distribution d'Airy. Le barycentre ne permet donc pas d'atteindre la borne de Cramér-Rao pour une distribution d'Airy, ni même non plus a priori pour les images tavelées à faible D/r_0 . Le résultat s'obtient pour le barycentre d'une gaussienne en dérivant la définition du maximum de vraisemblance avec l'expression analytique de la gaussienne, et il se déduit par exemple pour la distribution d'Airy de l'erreur infinie de l'estimation du barycentre dans ce cas-là.

5.1.4 Estimation optimale du déplacement à partir des images tavelées

5.1.4.1 Borne de Cramér-Rao pour l'estimation du centre des images tavelées sans connaissance de la phase

Il n'existe pas de relation analytique de la répartition spatiale d'énergie des images tavelées, qui est de nature fluctuante, de sorte qu'on ne peut pas calculer directement la borne de Cramér-Rao pour l'estimation de leur déplacement. Néanmoins en considérant la turbulence comme un bruit dégradant la distribution d'énergie d'Airy, on s'attend à ce que la borne de l'écart-type de leur déplacement soit de l'ordre de

$$\sigma_{\theta_k}^+ \simeq \frac{\kappa \times \sigma_{\text{tav}_k}}{\sqrt{N_{\text{ph}}}}, \quad (5.18)$$

où κ est un coefficient scalaire de l'ordre de 1, et σ_{tav_k} est l'écart-type de la distribution moyenne d'intensité des images courtes poses détiltées. Cette dernière distribution s'obtient en déconvoluant la distribution des images longues poses par celle du tip-tilt. Grâce à l'expression analytique (5.4) de l'écart-type du tip-tilt pour une turbulence de type Kolmogorov et en approximant les deux distributions par des gaussiennes on

obtient

$$\begin{aligned}\sigma_{\text{tav}_k} &= \sqrt{\sigma_{\text{pose}_k}^{\text{longue}}^2 - \sigma_{\theta_k}^{\text{temp}}^2} \\ &= 0.44 \times \lambda/r_0 \times \sqrt{1 - 0.942 \times (D/r_0)^{-1/3}},\end{aligned}\quad (5.19)$$

qui conduit finalement à

$$\sigma_{\theta_k}^+ \simeq \frac{0.44 \kappa \times \lambda/r_0}{\sqrt{N_{\text{ph}}}} \times \sqrt{1 - 0.942 \times (D/r_0)^{-1/3}}. \quad (5.20)$$

Comme la distribution moyenne d'intensité des images courtes poses détiltées est plutôt de type gaussien que de type Airy pour les forts D/r_0 , on peut supposer d'après la borne de Cramér-Rao pour une gaussienne à l'équation (5.9) que $\kappa \simeq 0.7$ dans ces cas-là. Quand le D/r_0 est faible au contraire, on s'attend d'après la borne de Cramér-Rao pour la distribution d'Airy à ce que κ soit légèrement inférieur.

5.1.4.2 Bénéfice attendu de la connaissance de la phase

Comme la structure tavelée des images est causée par les aberrations de phase hors-tilt uniquement, l'erreur minimum de centrage redevient de l'ordre de

$$\sigma_{\theta_k}^+ \simeq \frac{\kappa \times \lambda/D}{\sqrt{N_{\text{ph}}}}, \quad (5.21)$$

lorsque la phase détiltée φ_{det} est conjointement utilisée avec les mesures d'intensité de l'image dans l'expression (5.10) de la borne de Cramér-Rao. Puisque les images tavelées sont une version bruitée de la distribution d'Airy, on voit en effet que la connaissance de la phase permet de débruiter les images et donc de pouvoir considérer la distribution d'Airy pour la borne de Cramér-Rao. On obtient donc l'amélioration de précision rms de centrage suivante :

$$\rho(D/r_0) \simeq [0.44 \times D/r_0] \times \sqrt{1 - 0.942 \times (D/r_0)^{-1/3}}, \quad (5.22)$$

par rapport au cas où la phase n'est pas conjointement estimée, ce qui est représenté sur la figure 5.2.

J'ai vérifié ce bénéfice par des simulations, en ajoutant une contrainte supplémentaire de bruit de lecture, pour se placer dans un cas plus réaliste.

5.1.5 Simulations

J'ai considéré des D/r_0 allant de 2 à 30 (cf figure 5.3) et des précisions rms de la phase allant de 0 pour une phase parfaitement connue, à 1 radian, et cela indépendamment du bruit contenu dans les images. Les images tavelées sont simulées grâce à

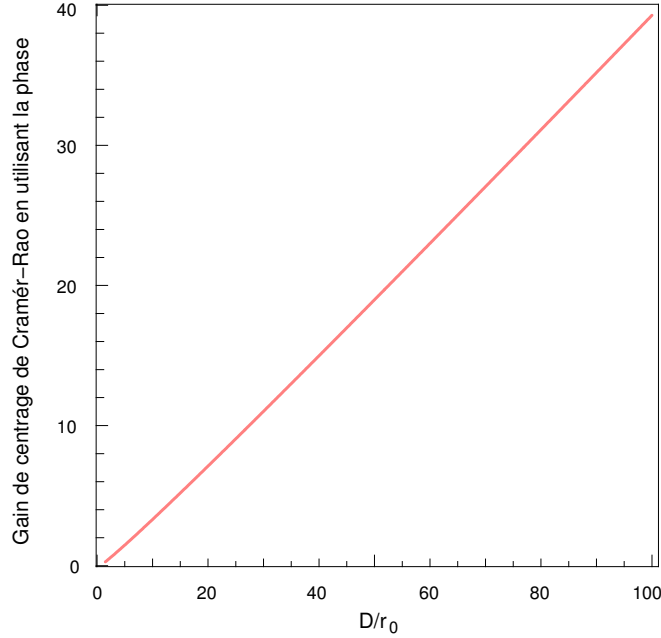


FIG. 5.2 – Gain de Cramér-Rao sur le centrage des images tavelées, apporté par la connaissance de la phase, en fonction du D/r_0 .

des écrans de phase Kolmogorov, comme cela est décrit dans le chapitre 3. La même définition du rapport signal à bruit par tavelure est également utilisée pour comparer les simulations en fonction du rapport D/r_0 :

$$\text{SNR} = \frac{N_{\text{ph}}/N_{\text{sp}}}{\sqrt{N_{\text{ph}}/N_{\text{sp}} + \sigma_{\text{CCD}}^2 \times s_{\text{sp}}^2}}, \quad (5.23)$$

où N_{ph} est le nombre moyen de photons dans l'image, $N_{\text{sp}} \sim (D/r_0)^2$ est le nombre de tavelures, et s_{sp} est la taille moyenne d'un speckle. Le rapport signal à bruit est fixé à ~ 10 pour l'ensemble des simulations, avec $\sigma_{\text{CCD}} = 0.25$ ph/pixel. Ce rapport signal à bruit correspond par exemple à $N_{\text{ph}} = 400$ photons pour $D/r_0 = 2$, et à $N_{\text{ph}} = 90\,100$ photons pour $D/r_0 = 30$.

5.1.5.1 Barycentre

Comme la fluctuation du nombre de speckles et de leur forme est plus significative pour les plus faibles (D/r_0), et que l'enveloppe des images tavelées n'est pas gaussienne pour les plus faibles (D/r_0), on s'attend à ce que l'erreur d'estimation du barycentre diminue avec l'augmentation du D/r_0 lorsque le RSB de l'équation (5.23) est constant. Cependant cette amélioration doit tendre à s'annuler asymptotiquement pour les forts D/r_0 .

On s'attend également à ce que l'erreur du barycentre diminue avec l'augmentation du D/r_0 à cause de la combinaison du bruit de lecture et de l'échantillonnage. En effet les taches images ne sont pas concentrées sur le même nombre de pixels (elles occupent de plus en plus de pixels avec l'augmentation du D/r_0), parce que l'échantillonnage des pixels a été pris constant dans toute la simulation avec $\lambda/D \sim 2$ pixels, pour l'estimateur de centre à phase connue. Les taches images correspondant aux faibles D/r_0 sont donc inutilement diluées dans le bruit de lecture.

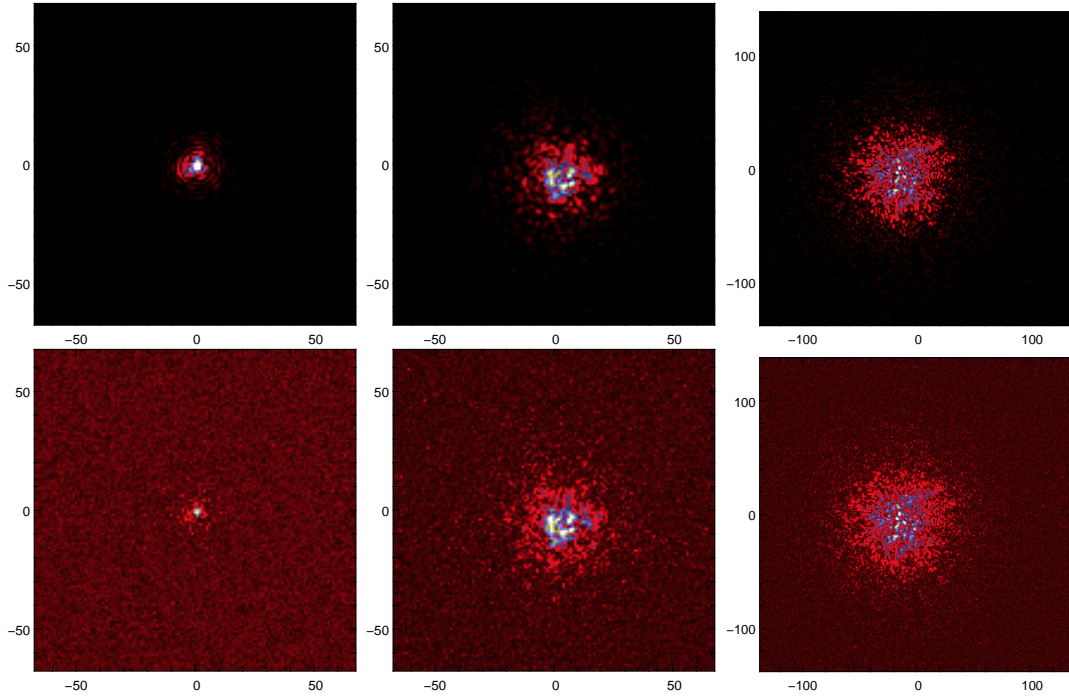


FIG. 5.3 – Illustration de la diversité des images simulées. Les images sans bruit sont représentées en haut, tandis que les images du bas sont les versions bruitées à partir desquelles le centrage est effectué (la phase est ici parfaitement connue). De gauche à droite les D/r_0 simulés sont de 2, 10 et 30.

Pour limiter cet effet pour les faibles D/r_0 , j'ai utilisé un estimateur amélioré par rapport au barycentre, qui utilise une fenêtre de centrage réduite par rapport à la taille de l'image, et qui me sert pour le calcul du barycentre dans les images bruitées. Notons au passage que le barycentre simulé est quand à lui déterminé par le simple barycentre de l'image non bruitée, et que je me suis d'ailleurs assuré de toujours conserver un champ suffisamment large pour ne pas introduire d'erreur de troncature dans le barycentre simulé. Le principe du barycentre amélioré est le suivant. La taille de la fenêtre de centrage est tout d'abord optimisée pour chaque D/r_0 considéré, en choisissant la troncature qui donne la plus faible erreur par rapport au barycentre simulé. De plus, j'utilise un algorithme itératif (conçu par Eric Thiébaud) pour réduire parallèlement le

biais et l'erreur causés par la troncature dissymétrique du signal par la fenêtre de centrage. L'algorithme met à jour à chaque itération et jusqu'à convergence, la position de la fenêtre de centrage circulaire apodisée avec le centre calculé à l'itération précédente. Comme on le voit sur la figure 5.4, ce barycentre amélioré permet de limiter les erreurs du barycentre liées à la combinaison du bruit de lecture et de l'échantillonnage, pour les faibles ($D/r_0 < 14$). En pratique le gain du barycentre amélioré dépend du niveau

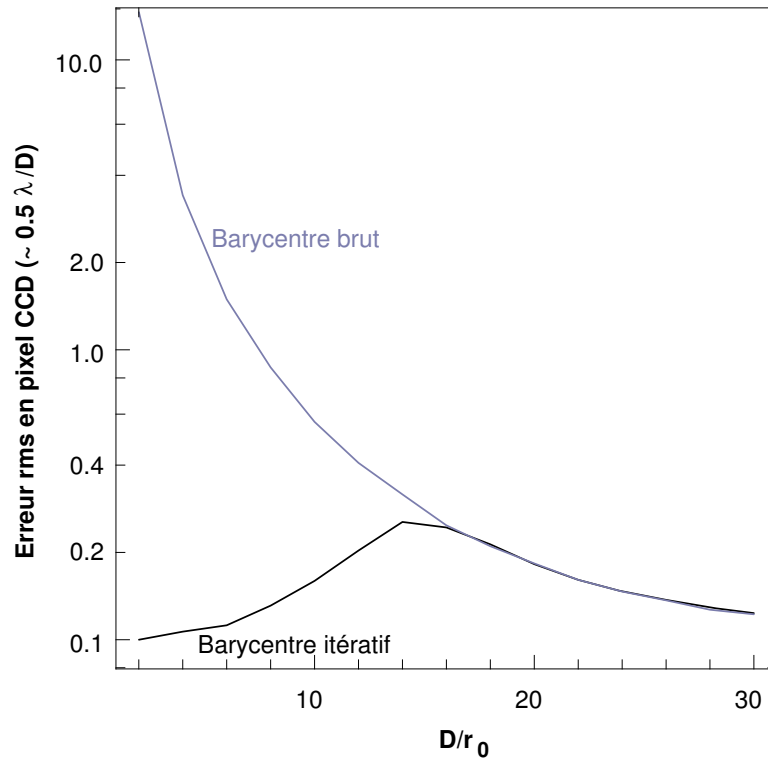


FIG. 5.4 – Performances du barycentre pour un RSB de 10, en fonction du D/r_0 et des conditions particulières d'échantillonnage par rapport au bruit de lecture.

de bruit de lecture, si bien que cet estimateur peut également se révéler utile pour des D/r_0 plus élevés, si le bruit de lecture est plus important. D'autres méthodes peuvent être envisagées et éventuellement combinées à celle-ci pour améliorer le barycentre, comme par exemple de ne prendre en compte qu'un nombre limité de pixels (les N plus brillants), ou alors tous ceux qui sont situés au dessus d'un certain seuil, ce qui n'a pas été mis en œuvre pour ces simulations.

5.1.5.2 Centrage à phase connue

En faisant l'hypothèse de bruit gaussien pour la combinaison du bruit de Poisson et du bruit du CCD, l'estimateur de maximum de vraisemblance du tip-tilt permettant

d'approcher la borne de l'équation (5.21) s'exprime par

$$\begin{aligned}\theta^+ &= \arg \min_{\theta} \left[(\mathbf{d} - \mathbf{m}(\theta | \varphi_{\text{det}}))^T \cdot \mathbf{C}_n^{-1} \cdot (\mathbf{d} - \mathbf{m}(\theta | \varphi_{\text{det}})) \right] \\ &= \arg \min_{\theta} f_{\text{centre}}(\theta),\end{aligned}\quad (5.24)$$

où \mathbf{d} est la répartition d'énergie détectée, $\mathbf{m}(\theta=0 | \varphi_{\text{det}})$ est le modèle centré de cette répartition étant donnée la phase détiltée φ_{det} , et $\mathbf{C}_n = \langle \mathbf{n} \cdot \mathbf{n}^T \rangle$ est la matrice de covariance du bruit et des erreurs de modélisation.

Le critère $f_{\text{centre}}(\theta)$ de l'équation (5.24) qu'il faut optimiser, est non-linéaire par rapport à θ . De plus on imagine facilement que ce critère puisse être multi-modal, en introduisant un décalage adapté dans la superposition des figures de speckles, c'est-à-dire un décalage tel que plusieurs speckles soient mis en coïncidence entre le modèle et les mesures, et aussi à cause du bruit. Le caractère multi-modal du critère s'observe effectivement dans la représentation bidimensionnelle des valeurs de $f_{\text{centre}}(\theta)$ sur la figure 5.5 (pour un $D/r_0 \sim 30$ et une précision rms sur la phase de 0.4 radian par

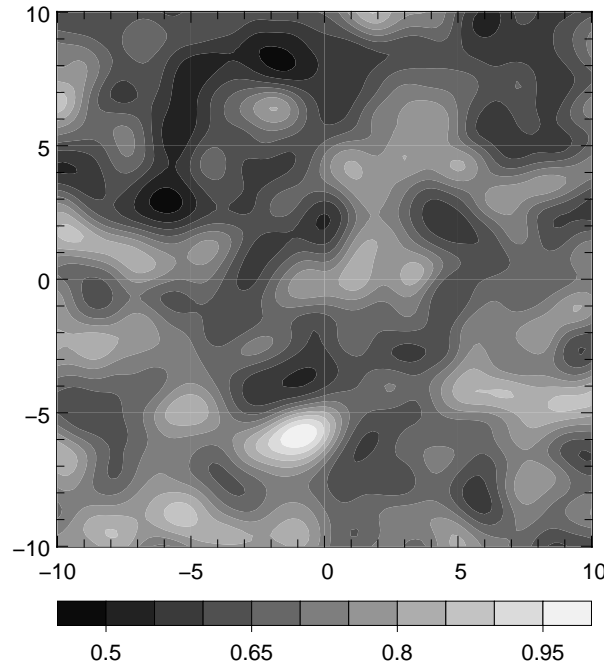


FIG. 5.5 – Représentation bidimensionnelle (en unité de pixel CCD sur les axes) des valeurs du critère $f_{\text{centre}}(\theta)$. On observe de nombreux optima locaux. Le meilleur centre obtenu est le minimum global de l'image.

exemple).

Un algorithme d'optimisation locale ne suffit donc pas pour optimiser $f_{\text{centre}}(\theta)$. Comme le nombre de paramètres est limité à 2, la stratégie d'optimisation globale que

j'ai employée est la suivante. J'ai calculé dans un premier temps le critère f_{centre} pour des valeurs de θ contenues dans une grille centrée sur une estimation initiale du barycentre et de taille et de nombre de points prédéfinis. Cette grille est suffisamment large par rapport à l'erreur d'estimation du tip-tilt par le barycentre mais également suffisamment fine par rapport à la taille typique des dépressions associées aux optima locaux. La position de la grille qui rend le critère f_{centre} minimum est ensuite raffinée à l'aide d'un algorithme d'optimisation locale de quasi-Newton, semblable à celui décrit à la section 3.5. L'erreur de centrage obtenue en fonction du D/r_0 et pour diverses précisions de connaissance de la phase est représentée sur la figure 5.6. On observe que

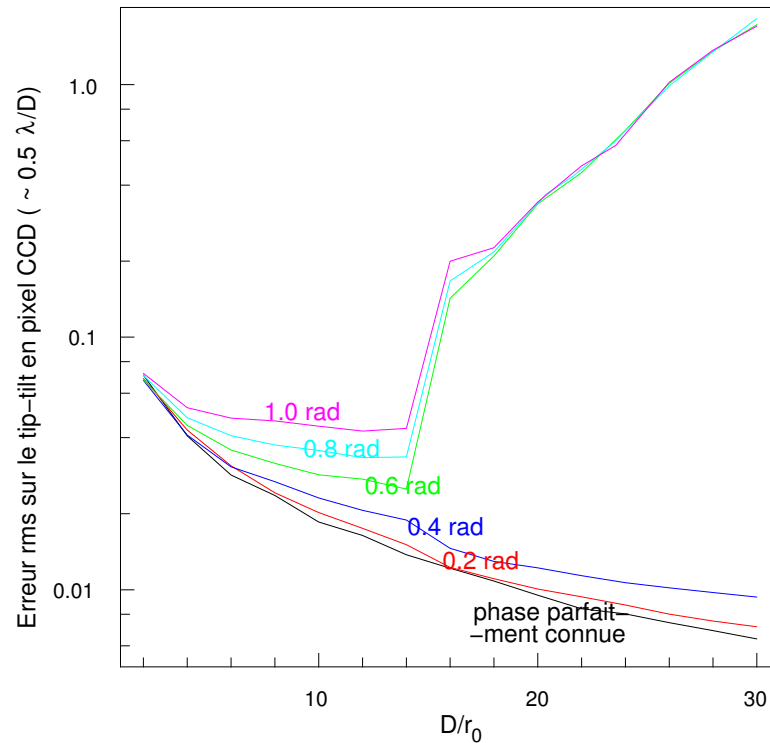


FIG. 5.6 — Performances de l'estimateur du tip-tilt utilisant la connaissance de la phase en fonction de D/r_0 pour différentes précisions de la phase φ_{det} .

l'optimisation globale devient moins efficace à partir de $D/r_0 \sim 16$, lorsque la phase est connue avec une précision supérieure à 0.6 radian et pour le niveau particulier de bruit considéré dans l'image, à cause de l'augmentation du nombre de minima locaux de f_{centre} avec le D/r_0 et avec le bruit. La stratégie d'optimisation globale peut alors être affinée en diminuant la taille de la grille utilisée lors de la première phase de l'optimisation. Toutefois j'ai vérifié en raffinant la grille que le problème vient du fait que l'optimum global de f_{centre} n'est alors plus un bon estimateur du tip-tilt. On peut recouper ce dysfonctionnement avec l'augmentation de la difficulté de la reconstruction

de phase avec le D/r_0 lorsque le tip-tilt n'est pas estimé suffisamment précisément à l'initialisation.

5.1.5.3 Gain de précision pour le centrage

Le gain de précision apporté par la connaissance de la phase dans les simulations que j'ai effectuées, n'est pas comparable au gain théorique de Cramér-Rao de l'équation (5.22), à cause de la présence du bruit de lecture, des conditions inadéquates d'échantillonnage pour le barycentre, et parce que le barycentre n'est pas non plus un estimateur de maximum de vraisemblance du centre des images tavelées pour des D/r_0 modérés. Le gain est donc en pratique plus important que sur la courbe 5.2, bien que le centrage à phase connue ait été effectué avec une pondération uniforme au lieu de la pondération exacte de bruit. Son évolution est également moins continue avec le D/r_0 à cause de l'impact de l'échantillonnage sur le barycentre. Par exemple le gain obtenu par rapport au barycentre pour $D/r_0=30$ est de 19.3, alors que le gain de Cramér-Rao qu'on obtient en théorie sans bruit de lecture est de 11. Je n'ai donc pas tracé la courbe de gain en fonction du D/r_0 . La validation de ce gain peut toutefois être estimée à partir des courbes de la figure 5.7.

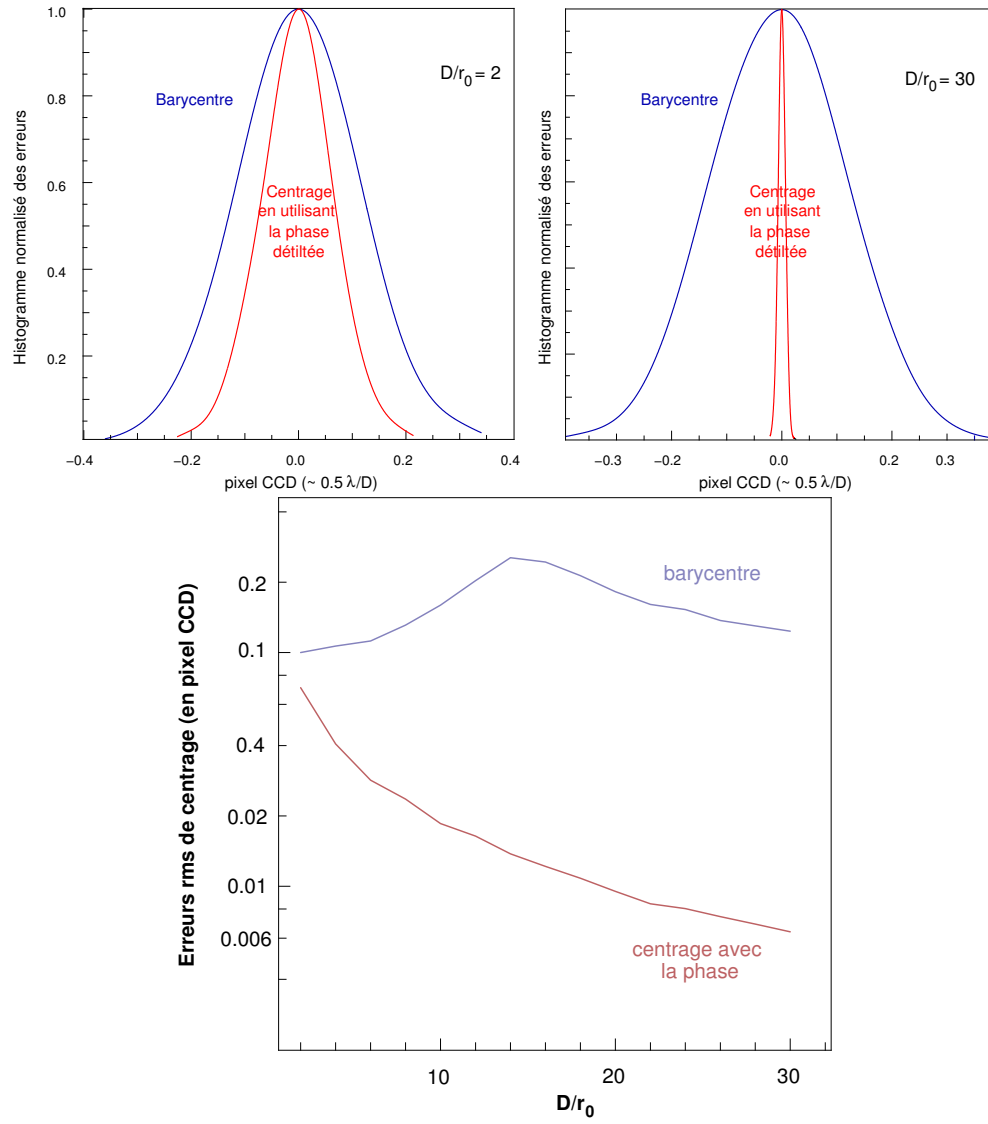


FIG. 5.7 – Gain de l'estimateur de centrage utilisant la connaissance de la phase par rapport au barycentre.

5.2 Estimation de la pente différentielle par interspectre bichromatique

L'estimation conjointe de la phase turbulente sur la pupille permet d'estimer très précisément le déplacement angulaire d'une source lumineuse à partir des images tavelées comme on vient de le voir, dans le cas où la source n'est pas résolue par le télescope. Cette estimation peut également être mise en oeuvre avec une source résolue, si la distribution angulaire d'énergie de la source peut être conjointement estimée, par déconvolution aveugle par exemple. Etant donné le gain en précision par rapport au barycentre mis en évidence à la section précédente, ces méthodes de centrage constituent un tout premier choix pour un projet comme ELP-OA, où il faut estimer le déplacement différentiel chromatique avec une très grande précision (cf. section 1.4.4.3).

Néanmoins les hautes fréquences spatiales des images peuvent être excessivement atténuées par la distribution angulaire d'énergie de la source, ce qui rend alors ces méthodes impraticables. Cette distribution angulaire, pour ELP-OA, dépend principalement du diamètre du projecteur laser (dans le cas d'un projecteur monolithique plutôt que interférométrique) et des conditions particulières de turbulence du site d'observation. On montre en effet [Fried, 1966] que le diamètre de télescope permettant d'obtenir la meilleure résolution dans le plan image, à partir d'images courtes poses détiltées et non corrigées d'une source ponctuelle vaut $\sim 3 r_0$. En inversant les plans objet et image, on peut approximativement utiliser le même raisonnement dans l'autre sens de traversée de l'atmosphère, pour déterminer le diamètre de projecteur laser qui permet de focaliser au mieux l'étoile laser. Typiquement $r_0 \sim 5$ cm à 589 nm pour un site comme l'OHP où est réalisé le démonstrateur ELP-OA, ce qui conduit à un diamètre de projecteur laser de l'ordre de quinze centimètres et une tache de ~ 20 cm dans la mésosphère à 589 nm, alors que le télescope de 1.52 m à l'OHP résout une tache de 2.2 cm. Cela se traduit par la coupure des hautes fréquences de la turbulence dans les images de la source, et empêche de pouvoir retrouver les aberrations d'ordre élevé de la phase à partir des images de la source.

Pour estimer le déplacement différentiel chromatique de la source dans un tel cas de figure, j'ai développé une méthode d'interspectre bichromatique, qui vise à exploiter un nombre de fréquences spatiales intermédiaire entre le barycentre et la reconstruction de phase, et à être ainsi moins sensible au diamètre de la source artificielle.

5.2.1 Principe de la mesure

Comme je l'ai montré à la section 1.1.2.2, l'autocorrélation de l'amplitude complexe sur la pupille du télescope et l'image au foyer sont associées linéairement par transformée de Fourier :

$$I_\lambda(\mathbf{u}) = \text{cste} \times \int_{\mathcal{P}} P_\lambda(\mathbf{f}) P_\lambda^*(\mathbf{f} + \mathbf{u}) \times e^{\frac{2\pi}{\lambda} j [\delta_\lambda(\mathbf{f}) - \delta_\lambda(\mathbf{f} + \mathbf{u})]} d\mathbf{f}, \quad (5.25)$$

où $I_\lambda(\mathbf{u})$ est la transformée de Fourier de l'image $i_\lambda(\mathbf{x})$ à la fréquence spatiale \mathbf{u} . En séparant le chemin optique entre une composante de tip-tilt et une partie détiltée :

$$\delta_\lambda(\mathbf{u}) = \boldsymbol{\theta}_\lambda \cdot \mathbf{u} + \delta'_\lambda(\mathbf{u}), \quad (5.26)$$

et en notant $P_\lambda^\times(\mathbf{f}, \mathbf{u}) = P_\lambda(\mathbf{f})P_\lambda^*(\mathbf{f} + \mathbf{u})$ le produit des transmissions de la pupille aux fréquences \mathbf{f} et $\mathbf{f} + \mathbf{u}$, il vient

$$I_\lambda(\mathbf{u}) = \text{cste} \times e^{-\frac{2\pi}{\lambda} j \boldsymbol{\theta}_\lambda \cdot \mathbf{u}} \int_{\mathcal{P}} P_\lambda^\times(\mathbf{f}, \mathbf{u}) \times e^{\frac{2\pi}{\lambda} j [\delta'_\lambda(\mathbf{f}) - \delta'_\lambda(\mathbf{f} + \mathbf{u})]} d\mathbf{f}. \quad (5.27)$$

L'interspectre bichromatique, en fonction d'un facteur α d'éventuel rééchantillonnage de I_{λ_2} , s'écrit

$$\begin{aligned} \mathcal{B}_{\lambda_1, \lambda_2}(\mathbf{u}, \alpha) &\triangleq I_{\lambda_1}(\mathbf{u}) \times I_{\lambda_2}^*(\alpha \mathbf{u}) \\ &= \text{cste} \times e^{-2\pi j [\boldsymbol{\theta}_{\lambda_1} / \lambda_1 - \alpha \boldsymbol{\theta}_{\lambda_2} / \lambda_2] \cdot \mathbf{u}} \times \\ &\quad \iint_{\mathcal{P}_{1,2}} P_{\lambda_1, \lambda_2}^\times(\alpha, \mathbf{f}, \mathbf{u}) \times e^{2\pi j \left([\delta'_{\lambda_1}(\mathbf{f}) - \delta'_{\lambda_1}(\mathbf{f} + \mathbf{u})] / \lambda_1 - [\delta'_{\lambda_2}(\alpha \mathbf{f}') - \delta'_{\lambda_2}(\alpha (\mathbf{f}' + \mathbf{u}))] / \lambda_2 \right)} d\mathbf{f} d\mathbf{f}'. \end{aligned} \quad (5.28)$$

où $P_{\lambda_1, \lambda_2}^\times(\alpha, \mathbf{f}, \mathbf{u}) = P_{\lambda_1}^\times(\mathbf{f}, \mathbf{u}) P_{\lambda_2}^\times(\alpha \mathbf{f}, \alpha \mathbf{u})$.

Dans un domaine de fréquences spatiales proches de zéro, on s'attend à ce que le tip-tilt différentiel chromatique domine l'expression (5.28). Le principe de mesure de la pente différentielle par interspectre bichromatique consiste donc à ajuster l'angle du complexe $\mathcal{B}_{\lambda_1, \lambda_2}(\mathbf{u}, \alpha)$. Cela correspond également à ajuster l'angle minimisant la distance entre $I_{\lambda_1}(\mathbf{u})$ et $I_{\lambda_2}(\alpha \mathbf{u})$ au sens des moindres carrés. En effet on a

$$\begin{aligned} \boldsymbol{\Theta}^+ &= \arg \min_{\boldsymbol{\Theta}} \sum_{\mathbf{u}} w(\mathbf{u}) \times |d_1(\mathbf{u}) - e^{2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} d_2(\alpha \mathbf{u})|^2 \\ &= \arg \min_{\boldsymbol{\Theta}} \sum_{\mathbf{u}} w(\mathbf{u}) \times \left(\text{Re}^2 \left[d_1(\mathbf{u}) - e^{2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} d_2(\alpha \mathbf{u}) \right] + \text{Im}^2 \left[d_1(\mathbf{u}) - e^{2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} d_2(\alpha \mathbf{u}) \right] \right) \\ &= \arg \min_{\boldsymbol{\Theta}} \sum_{\mathbf{u}} -2w(\mathbf{u}) \times \left(\text{Re}[d_1(\mathbf{u})] \times \text{Re}[e^{2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} d_2(\alpha \mathbf{u})] + \text{Im}[d_1(\mathbf{u})] \times \text{Im}[e^{2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} d_2(\alpha \mathbf{u})] \right) \\ &= \arg \min_{\boldsymbol{\Theta}} \sum_{\mathbf{u}} -w(\mathbf{u}) \times \text{Re} \left[2 d_1(\mathbf{u}) d_2^*(\alpha \mathbf{u}) e^{-2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} \right] \\ &= \arg \min_{\boldsymbol{\Theta}} \sum_{\mathbf{u}} w(\mathbf{u}) \times \text{Im}^2 \left[d_1(\mathbf{u}) d_2^*(\alpha \mathbf{u}) e^{-2\pi j \boldsymbol{\Theta} \cdot \mathbf{u}} \right], \end{aligned} \quad (5.29)$$

où $d_k(\mathbf{u}) = I_{\lambda_k}(\mathbf{u}) / |I_{\lambda_k}(\mathbf{u})|$, et en utilisant la formule trigonométrique $\text{Im}^2[x] = 1 - \text{Re}[2x]$ pour obtenir la dernière égalité. C'est cette dernière égalité, qui optimise au sens des moindres carrés l'arc moyen et donc l'angle moyen du complexe $\mathcal{B}_{\lambda_1, \lambda_2}(\mathbf{u}, \alpha)$, que j'utilise finalement pour estimer le déplacement différentiel chromatique de la source.

Comme la première égalité de l'équation (5.29) correspond à optimiser le déplacement différentiel des images par un maximum de corrélation entre les deux images, au

rééchantillonnage et à la pondération $w(\mathbf{u})$ près, on voit que ma mesure peut être considérée comme une version améliorée d'un maximum de corrélation, avec une possible mise à l'échelle de la taille d'une des images et une pondération adaptée des fréquences spatiales en fonction de la ressemblance effective des images à ces fréquences.

Le poids $w(\mathbf{u})$ est obtenu empiriquement par

$$w(\mathbf{u}) = \left\langle \text{Im}^2 \left(\mathcal{B}_{\lambda_1, \lambda_2}^0(\mathbf{u}, \alpha) \right) \right\rangle^{-1}, \quad (5.30)$$

où $\mathcal{B}_{\lambda_1, \lambda_2}^0(\mathbf{u}, \alpha)$ est l'interspectre des images détiltées et normalisées, et où le moyennage est effectué sur un nombre significatif de réalisations indépendantes. En pratique, pour des données non simulées, on ne dispose pas des images recentrées pour calculer $w(\mathbf{u})$. On peut dans ce cas-là mettre en oeuvre une auto-calibration, en calculant de façon itérative un poids $w(\mathbf{u})^{(n)}$, qui bénéficie de l'amélioration progressive de recentrage des images au fur et à mesure de l'amélioration de l'estimation de la pente dans l'interspectre bichromatique.

5.2.2 Simulations

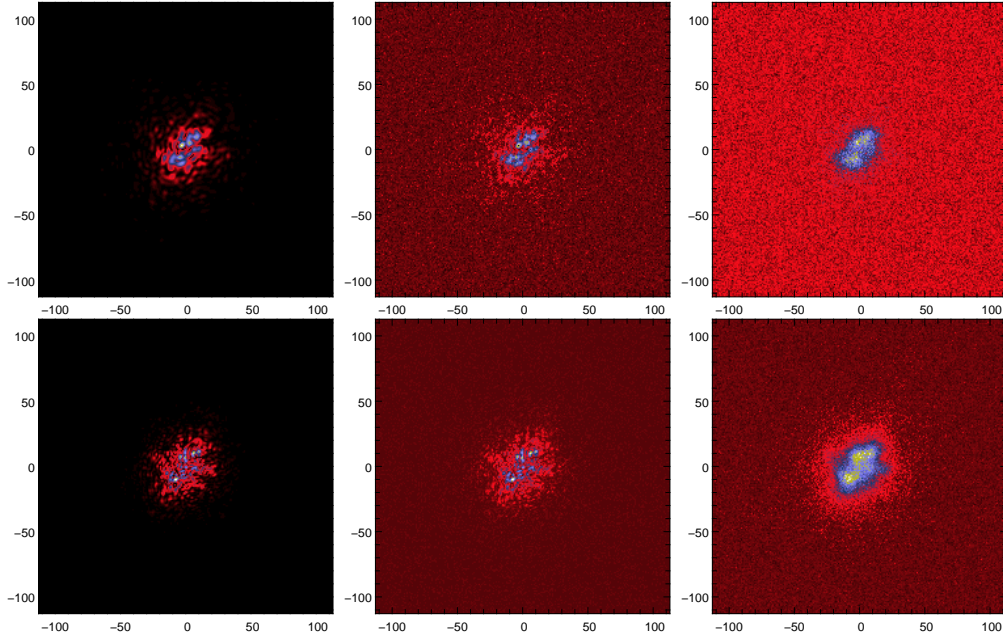


FIG. 5.8 — Images typiques simulées pour l'interspectre bichromatique. De gauche à droite sont représentées l'image sans bruit d'une source ponctuelle, l'image bruitée et enfin l'image bruitée d'une source résolue de largeur à mi-hauteur λ_1/r_0 en unité angulaire. Les images sont représentées pour les deux longueurs d'onde λ_1 et λ_2 , en haut et en bas respectivement, et sans rééchantillonnage ($\alpha=1$).

Mon objectif étant avant tout de débroussailler une nouvelle méthode, j'ai effectué des simulations pour un seul couple de longueurs d'onde $[\lambda_1, \lambda_2] = [569, 330]$ nm, une condition particulière de turbulence $D/r_0(330 \text{ nm}) = 18$, et un rapport signal à bruit — tel que défini par l'équation (5.23) — particulier de 10.5 pour les deux longueurs d'onde. Ce RSB correspond à un nombre moyen de 36 000 et 8 300 photons pour 330 et 569 nm respectivement, et à un bruit de détecteur de $\sigma_{\text{CCD}} = 0.25$ photons/pixel. On peut s'attendre à une certaine dépendance des performances de la méthode avec le couple de longueurs d'onde, avec la force de la turbulence, avec la balance de RSB entre les deux longueurs d'onde, avec le niveau relatif du bruit de détecteur dans le bruit total, etc... Une étude exhaustive de la méthode resterait donc à faire en fonction de tous ces paramètres. Je me suis ici uniquement intéressé au facteur α de rééchantillonnage, à la portion utile de l'interspectre, et à l'extension de la source laser artificielle.

Un jeu d'images typiquement simulées est représenté sur la figure 5.8, pour une source lumineuse à 589 nm non résolue ou de diamètre $\lambda_1/r_0(\lambda_1)$.

La comparaison des erreurs d'estimation du déplacement différentiel chromatique, entre le barycentre et l'interspectre bichromatique, est représentée sur la figure 5.9 en écart-type, et en fonction du facteur relatif α de rééchantillonnage de la pupille.

La valeur $\alpha = 1$ correspond au rééchantillonnage optique de la pupille de la longueur d'onde λ_1 vers λ_2 , par le facteur λ_1/λ_2 "naturel". C'est en effet le rééchantillonnage qu'il convient d'effectuer sur la phase en l'absence de grandissement optique dans le montage pour obtenir les images dans le plan de Fourier conjugué. La courbe 5.9 correspond à une source ponctuelle, et l'erreur sur le déplacement différentiel a été ramenée a posteriori à l'erreur sur le déplacement à la longueur d'onde λ_1 , afin de comparer les erreurs d'une grandeur invariante avec le facteur de rééchantillonnage α . On observe un gain optimum de 4.1 entre l'interspectre et le barycentre pour $\alpha^+ \simeq (\lambda_2/\lambda_1)^{0.2}$. Cette valeur correspond à corrélérer des images ayant des tailles d'enveloppes similaires, ce qui traduit bien que les hautes fréquences spatiales ne sont pas utilisées. C'est d'ailleurs pour ce facteur de rééchantillonnage optimum α^+ que les poids utiles $w(\mathbf{u})$ de l'interspectre peuvent être exploités au plus loin, comme on le voit sur la figure 5.10 représentant la coupe radiale des poids $w(\mathbf{u})$ de l'interspectre en fonction du facteur de rééchantillonnage α .

Les résultats montrent en définitive que la mesure de déplacement par interspectre est adaptée pour l'exploitation des basses fréquences principalement, à défaut d'exploiter la structure tavelée des images par de la reconstruction de phase.

Il faut noter que le barycentre est légèrement défavorisé dans cette comparaison par l'échantillonnage trop important des images. Toutefois le bruit de lecture est très inférieur au bruit de photons, et on s'attend donc à ce que le gain significatif obtenu pour le meilleur facteur de rééchantillonnage α^+ soit en partie conservé avec un échantillonnage optimal des images pour l'estimation du barycentre — lorsque le bruit de photons domine toutefois.

J'ai dans un second temps effectué une simulation avec une source étendue de

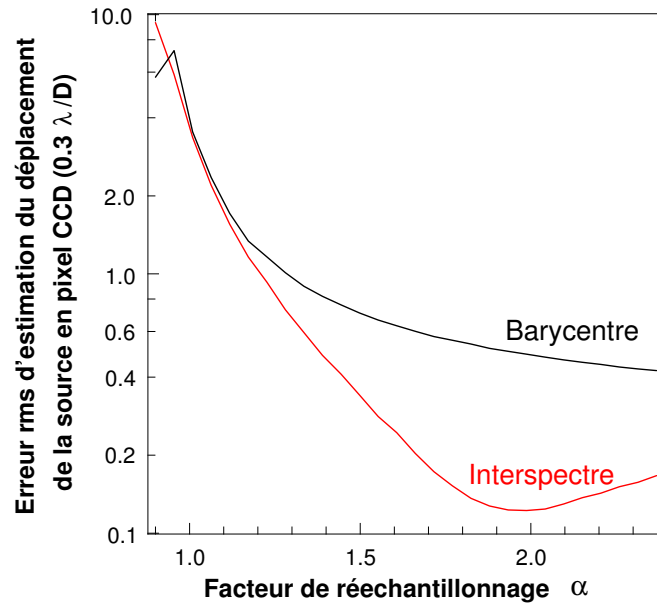


FIG. 5.9 – Gain de centrage de l'interspectre bichromatique par rapport au barycentre, en fonction du facteur de rééchantillonnage α .

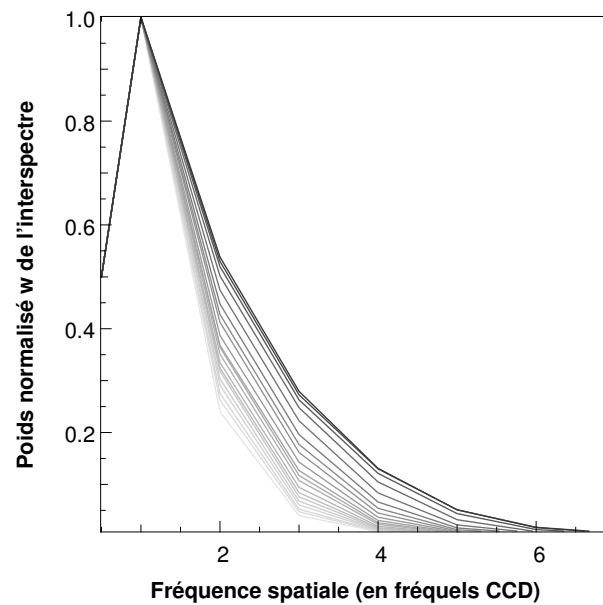


FIG. 5.10 – Coupe radiale des poids de l'interspectre en fonction de la fréquence spatiale. Les différents facteurs de rééchantillonnage α sont représentés par une couleur foncée croissant avec α . L'augmentation de α est stoppée à son optimum $\alpha^+ \simeq (\lambda_2/\lambda_1)^{0.2}$ pour une meilleure visualisation.

diamètre λ_1/r_0 , pour estimer la perte de précision causée par une coupure brutale des hautes fréquences spatiales des images. J'ai pris directement le meilleur facteur d'échantillonnage α^+ obtenu pour une source ponctuelle, bien que celui-ci puisse de nouveau être réoptimisé pour les images de la source résolue. En effet j'ai supposé que la réoptimisation n'était pas critique, vu que ce sont principalement les basses fréquences qui sont exploitées dans cette mesure. Comme escompté je n'obtiens pas de perte significative du gain de l'interspectre bichromatique pour une source fortement résolue. En effet le gain de l'erreur de centrage rms par rapport au barycentre est de 3.95 au lieu de 4.1 pour la source ponctuelle.

5.3 Amélioration de la résolution spatiale d'un détecteur à comptage de photons

5.3.1 Présentation des caméras CPng/Algol

Les caméras CPng/Algol sont des CCD intensifiés de nouvelle génération, permettant de détecter des photons individuels dans le visible à la cadence de 265 images par seconde. Elles ont été assemblées par l'équipe AIRI et l'OCA à partir de composants commerciaux et bénéficient d'un traitement du signal temps réel évolué, au développement duquel j'ai participé, pour améliorer la détection et le centrage des photoévénements. Elles font partie d'une collaboration pluridisciplinaire entre le CRAL et l'OCA en astronomie et le Centre Commun de Quantimétrie en biologie. En Astronomie les caméras sont utilisées pour l'interférométrie des tavelures et pour l'interférométrie à plusieurs télescopes ; en Biologie elles sont utilisées pour caractériser l'évolution des cellules par bioluminescence. D'autres caméras, comme les EMCCD par exemple, peuvent être utilisées pour les acquisitions à bas flux et à cadence élevée dans le visible. Mais ces-dernières sont encore pour le moment trop bruyantes, pour la détection de photons individuels à la cadence rapide de CPng, et sauf à regrouper fortement leurs pixels, les EMCCD sont plus lentes.

La conception matérielle de la caméra est représentée sur la figure 5.11. Le principe optoélectronique de base est le suivant. Un photon incident est détecté par une photocathode semi-conductrice avec une probabilité de 30 à 40 % d'absorption et de libération d'un électron. Cet électron est ensuite accéléré et multiplié lors de la traversée du canal le plus proche d'une galette de micro-canaux sous haute tension. L'intérieur de chaque canal est en effet spécialement traité pour favoriser l'émission secondaire d'électrons. Le gain de multiplication d'une galette vaut entre 10^3 et 10^5 typiquement. Après multiplication, le faisceau d'électrons est converti à l'aide d'un phosphore en faisceau de photons, lequel va finalement illuminer un CCD pour la lecture. Le but de la multiplication est que le faisceau de photons correspondant à un unique photon incident, soit facilement détectable dans l'image CCD par rapport au niveau de bruit du CCD. L'en-

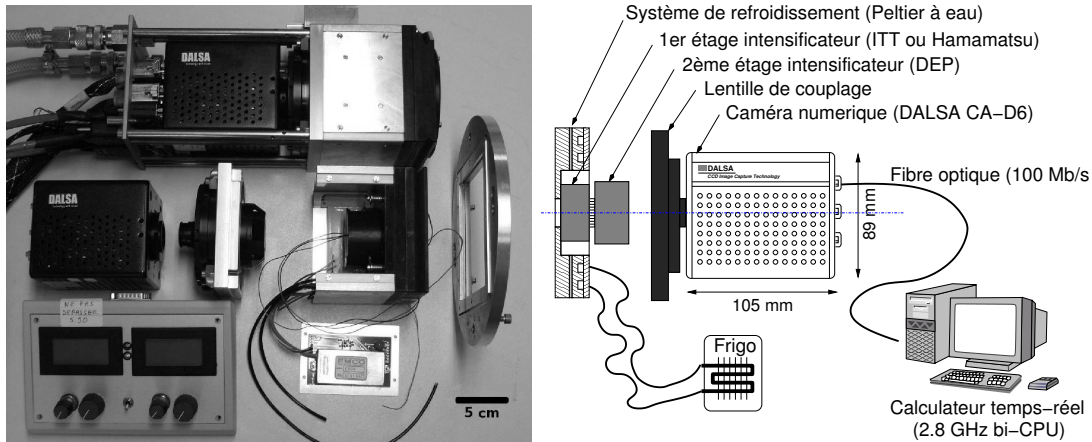


FIG. 5.11 – Conception matérielle des CCD intensifiés de nouvelle génération CPng

semble photocathode, galette de microcanaux et phosphore constitue un *intensificateur d'image*. Pour CPng deux intensificateurs ont été cascades afin d'obtenir un gain suffisant de $\sim 10^4$ – 10^5 par rapport au bruit de lecture du CCD, qui après débiaisage peut être considéré uniforme spatialement, avec un écart-type temporel pour chaque pixel qui vaut $\sigma_{\text{CCD}} = 0.5$ ADU. Le premier intensificateur a été choisi pour la sensibilité de sa photocathode, puisque celle-ci fixe le rendement quantique global de la caméra, avec la qualité de l'algorithme de détection. Pour couvrir toute la gamme du visible avec un rendement quantique comparable, deux caméras ont d'ailleurs été construites, une couvrant le domaine spectral bleue et l'autre le rouge, différant uniquement par la sensibilité chromatique de leur photocathode d'entrée.

Une lentille de focalisation est utilisée pour coupler le faisceau intensifié de photons et le CCD de lecture. L'idéal pour la détection est de focaliser chaque photoévenement sur un seul pixel CCD afin de maximiser le rapport signal à bruit. La focalisation a cependant été ajustée sur une largeur à mi-hauteur moyenne des photoévenements de 1.5 pixels CCD. En effet ce compromis permet d'effectuer un recentrage sous-pixel des positions des photoévenements, en plus d'une détection efficace, tout en limitant le trou de comptage de photons, c'est à dire le volume spatio-temporel dans lequel on ne peut détecter qu'un seul photon même s'il en arrive plusieurs (à cause du principe de comptage de photons).

5.3.2 Evènements parasites

Certains bruits sont inhérents aux intensificateurs et viennent occasionnellement parasiter le signal photonique. C'est le cas notamment des photoévenements thermiques, des ionoévenements et des évènements rémanents (et des rayons cosmiques

sur le ciel).

Les photoévénements thermiques sont liés au courant thermique des photocathodes et suivent donc la statistique de Fermi-Dirac. En pratique j'ai même retrouvé le *gap* ou zone d'énergie interdite du semi-conducteur à partir des mesures du courant thermique en fonction de la température, comme cela est expliqué plus en détail dans l'article CPng mis en annexe. Comme on le voit sur la figure 5.12, le courant thermique est réduit à moins de 0.5 électron par trame, en abaissant la température de la photocathode d'entrée jusqu'à -35°C à l'aide de Peltier à eau. J'ai également pu vérifier en

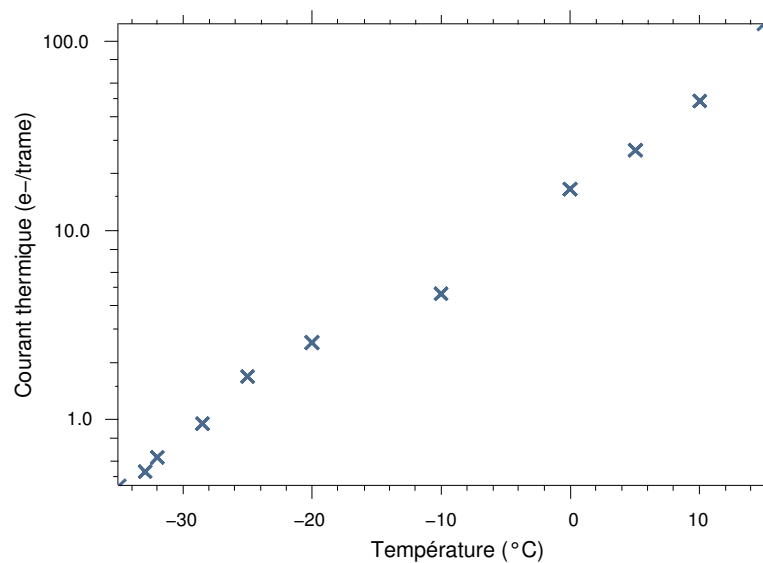


FIG. 5.12 – Mesure du courant thermique observé dans l'image en fonction de la température de refroidissement de la photocathode d'entrée. La température courante de fonctionnement est de -35°C

pratique en éteignant le premier intensificateur, que les électrons thermiques de la seconde photocathode ne sont pas suffisamment amplifiés pour être détectables sur le CCD.

Les ionoévénements sont créés par le détachement de cations au niveau des gallettes de microcanaux. Ces cations sont accélérés dans le sens inverse des électrons et viennent impacter la photocathode, ce qui a pour effet d'arracher un petit paquet d'électrons. Cela résulte après la retransmission des intensificateurs en un événement géant dans l'image CCD, comme on l'observe sur la figure 5.13. Le nombre typique d'occurrences de ionoévénements est de un toutes les 20 trames. En pratique les ionoévénements peuvent être facilement repérés et exclus du traitement photonique, grâce à leur extension spatiale particulière.

Les événements rémanents sont des événements détectés dans des trames successives, à cause de la rémanence des phosphores. De l'ordre de 25 % des événements

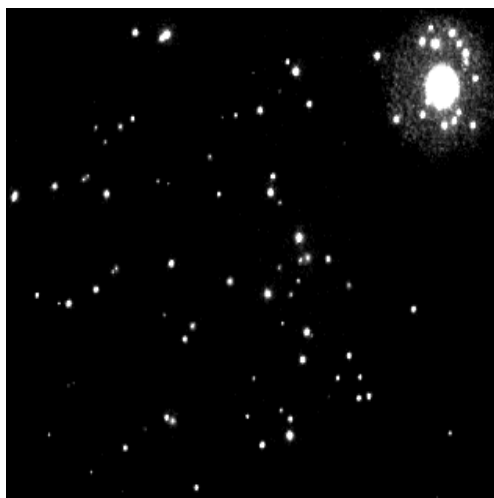


FIG. 5.13 – Ionoévènement parasite (en haut à droite) dans une trame CCD

sont rémanents d'une trame à l'autre, comme on le voit sur l'intercorrélacion spatio-temporelle des photoévènements détectés, représentée sur la figure 5.14. La rémanence

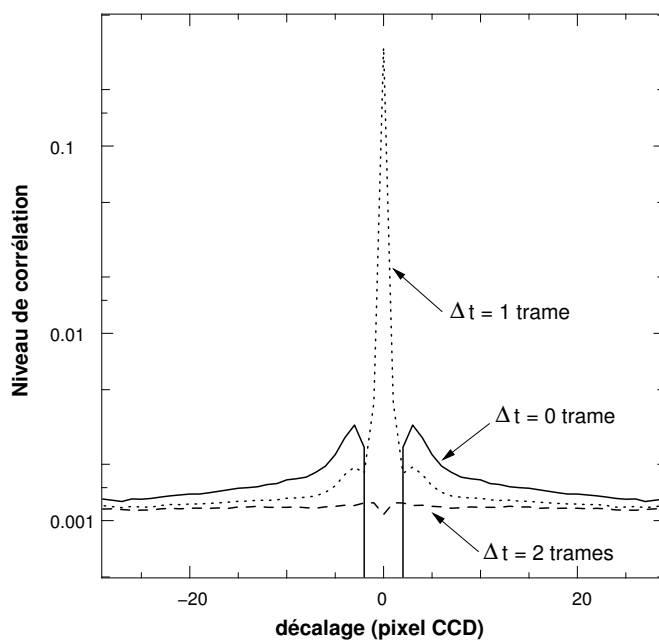


FIG. 5.14 – Intercorrélacion spatio-temporelle des photoévènements détectés. On observe le trou de comptage de photons pour $\Delta t = 0$ et le pic de rémanence pour $\Delta t = 1$. Il n'y a pas de rémanence pour $\Delta t = 2$.

pourrait toutefois être utilisée au même titre que l'extension spatiale des photoévénements pour effectuer un recentrage temporel des événements. Mais je n'ai pas eu le temps de mettre en pratique cette idée.

5.3.3 Modèle de maximum de vraisemblance pour la détection des photoévénements

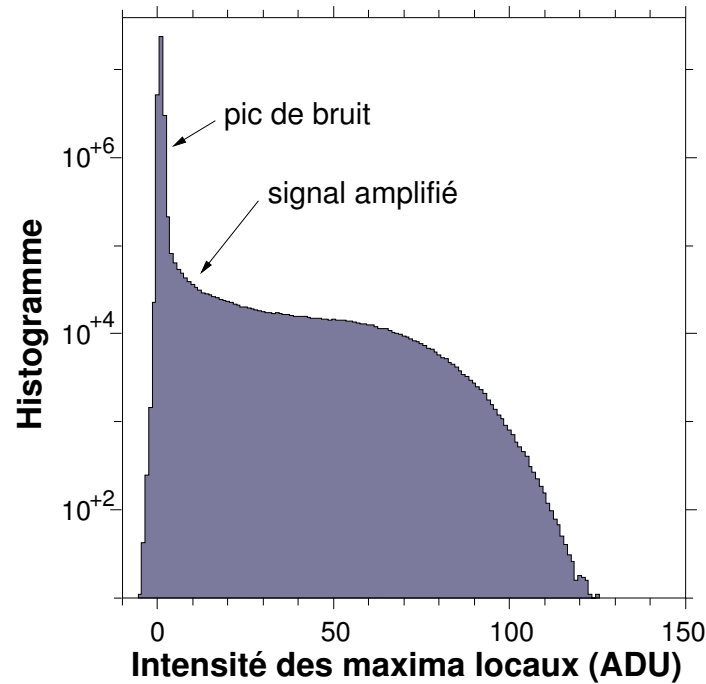


FIG. 5.15 – Histogramme des intensités des maxima locaux dans les images.

La distribution énergétique des maxima locaux détectés dans les images brutes, figure 5.15, montre que les photoévénements sont plus ou moins brillants par rapport au bruit du détecteur. La distribution gaussienne qu'on observe à gauche correspond à la distribution énergétique des pics locaux de bruit, tandis que la décroissance exponentielle qui suit correspond à la distribution énergétique des pics de photoévénements, et traduit la réponse énergétique des deux intensificateurs combinés. La détection doit donc permettre d'aller piocher dans le bruit un certain nombre d'événements peu brillants.

Comme on le voit sur la figure 5.16, les photoévénements ont une forme $s(\mathbf{x})$ en cloche, où \mathbf{x} est la position sur le CCD. J'ai également pu vérifier que l'écart-type de cette distribution est assez faible ; que la forme ne change pas selon que le photoévénement est plus ou moins brillant ; et que la forme est la même dans tout le champ

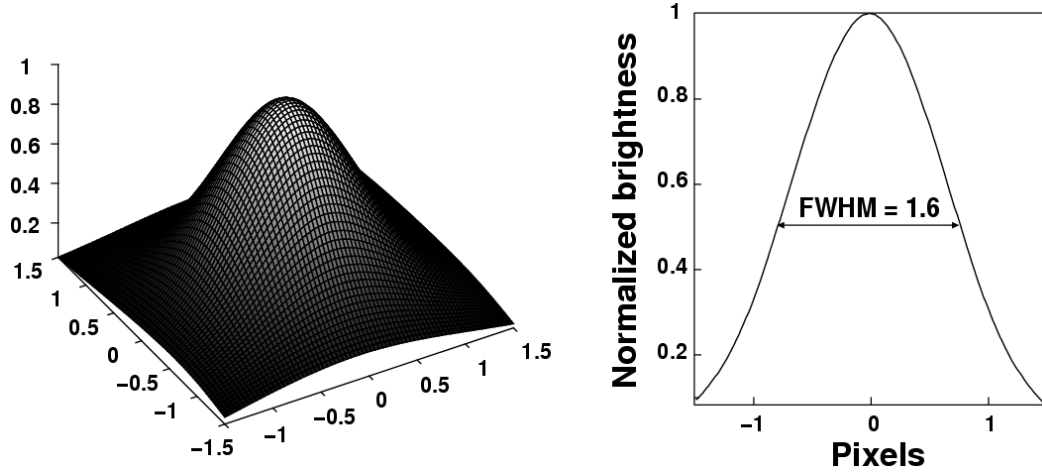


FIG. 5.16 – Distribution spatiale de l'énergie d'un photoévènement mesurée dans les images CCD et normalisée par rapport au pic du photoévènement.

du CCD, ce qui a tout de même nécessité d'ajuster la focalisation de la lentille. En pratique $s(\mathbf{x})$ est calibrée sur les données et est assez bien approximée par une fonction spline bi-cubique séparable de même largeur à mi-hauteur.

L'image débiaisée peut être modélisée par

$$d(\mathbf{x}) = m(\mathbf{x}) + n(\mathbf{x}), \quad (5.31)$$

où $m(\mathbf{x})$ modélise les photoévènements de l'image, et où $n(\mathbf{x})$ prend en compte le bruit du CCD et les fluctuations du modèle. En notant α_k l'intensité du pic du k -ième évènement et \mathbf{x}_k la position de ce pic, le modèle des photoévènements s'écrit

$$m(\mathbf{x}) = \sum_k \alpha_k s(\mathbf{x} - \mathbf{x}_k). \quad (5.32)$$

Sous l'hypothèse de bruit gaussien uniforme, la détection des photoévènements de l'image au sens du maximum de vraisemblance consiste à minimiser

$$\chi^2 = \frac{1}{\sigma_{\text{CCD}}^2} \sum_{\mathbf{x}} [d(\mathbf{x}) - m(\mathbf{x})]^2, \quad (5.33)$$

par rapport aux α_k et aux \mathbf{x}_k . On note qu'en plus d'être optimale, cette détection effectuée simultanément un recentrage automatique des photoévènements. De plus, on peut imaginer raffiner encore plus le modèle de détection en rajoutant la dépendance temporelle des photoévènements. Cela permettrait grâce à la rémanence d'effectuer un recentrage spatio-temporel automatique des photoévènements.

5.3.4 Détection temps réel

Malheureusement le critère de détection globale exprimé à l'équation (5.33) est aujourd'hui trop long pour pouvoir être optimisé en temps réel. En effet il nécessite de nombreuses itérations parce que le nombre de photoévénements dans l'image n'est pas connu au départ de l'optimisation. De plus la détection doit nécessairement être effectuée en temps réel pour réduire le volume de stockage des données.

En supposant que les événements ne se chevauchent pas trop, on simplifie donc le critère de détection globale par un critère de sélection locale d'un seul événement. On considère que la présence d'un événement indépendant des autres en \mathbf{x} doit minimiser

$$\chi_{\text{local}}^2(\alpha) = \frac{1}{\sigma_{\text{CCD}}^2} \sum_{\Delta \mathbf{x} \in S} [d(\mathbf{x} + \Delta \mathbf{x}) - \alpha s(\Delta \mathbf{x})]^2, \quad (5.34)$$

où S est le support spatial de la forme s des photoévénements. L'optimisation linéaire de χ^2 par rapport au paramètre α conduit à sélectionner les maxima locaux de l'image d_c corrélée par la forme s :

$$d_c(\mathbf{x}) = \sum_{\Delta \mathbf{x} \in S} d(\mathbf{x} + \Delta \mathbf{x}) s(\Delta \mathbf{x}). \quad (5.35)$$

Cette corrélation de l'image est plus connue sous le nom de filtrage adapté. Elle a pour effet de réhausser le niveau de certains photoévénements faibles par rapport au bruit. En effet le gain en rapport signal à bruit pour chaque photoévénement de l'image s'écrit :

$$\begin{aligned} \rho &= \frac{\langle d_c(\mathbf{x}_k) \rangle}{\langle d(\mathbf{x}_k) \rangle} \times \frac{\sigma_{d(\mathbf{x}_k)}}{\sigma_{d_c(\mathbf{x}_k)}} \\ &= \sqrt{\sum_{\Delta \mathbf{x} \in S} \left(\frac{s(\Delta \mathbf{x})}{s(\mathbf{0})} \right)^2} \\ &\simeq 1.5 \end{aligned} \quad (5.36)$$

Tous les maxima locaux de l'image corrélée d_c ne sont toutefois pas des pics de photoévénements, à cause du bruit, et il faut sélectionner les maxima locaux les plus brillants. En pratique cela est fait à l'aide d'un seuil de détection ϵ , choisi pour avoir une probabilité de détection de moins d'un pic de bruit par image :

$$\Pr [d(\mathbf{x}) \geq \epsilon] = \frac{1}{2} \operatorname{erfc} \left(\frac{\epsilon}{\sqrt{2} \sigma_{\text{CCD}}} \right) < \frac{1}{N_{\text{pix}}}, \quad (5.37)$$

où $\operatorname{erfc}(t) = \frac{2}{\sqrt{\pi}} \int_t^{+\infty} e^{-u^2} du$ est la fonction d'erreur complémentaire associée à la probabilité gaussienne, et N_{pix} est le nombre de pixels du CCD. Avec $\sigma_{\text{CCD}} = 0.5$ ADU on

obtient $\epsilon \geq 2.2$ ADU. Comme les photoévénements rémanents sont plus nombreux à faible intensité, on prend en pratique une valeur de seuil légèrement plus élevée (4 – 5 ADU). La valeur de seuil doit également être réhaussée pour le cas où la précision de centrage des photoévénements primerait sur la détection. En effet le centrage d'un photoévènement utilise les intensités des pixels dans une boîte 3×3 centrée sur le pic du photoévènement. Les intensités autour du pic doivent donc aussi avoir un niveau significatif par rapport au bruit, ce qui réhausse le niveau de seuil pour le pic. Au final la probabilité de détection d'un faux photoévènement dans l'image est négligeable.

En pratique j'ai appliqué la méthode de détection que je viens de décrire sur la mesure de rendement quantique des caméras CPng, en retirant du comptage les ionoévènements et statistiquement les photoévènements thermiques et rémanents. Les flux de référence utilisés pour la mesure sont délivrés par des photodiodes, calibrées à l'OCA par Alain Blazit, avec une précision de 2%. J'ajuste les points de mesure avec les courbes de rendement quantique des photocathodes d'entrée fournies par les constructeurs, en fonction d'un facteur de perte achromatique η_{perte} , puisque le signal est achromatique après la photocathode d'entrée. Le rendement quantique effectif des caméras CPng est représenté sur la figure 5.17. Il vaut entre 26 et 36 % dans tout le visible, et confirme le bénéfice attendu de la meilleure sensibilité des photocathodes et de la réduction de l'obstruction du signal par les micro-canaux, par rapport aux générations précédentes de caméra à comptage de photons (la dernière caméra de l'équipe CP40 avait un rendement de 15%).

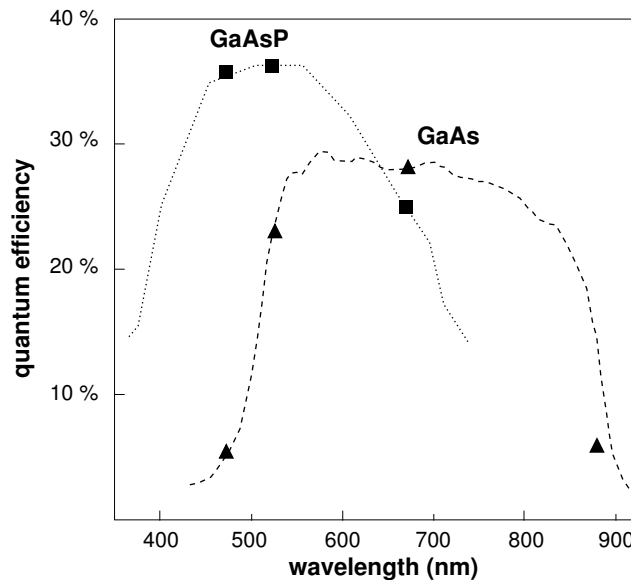


FIG. 5.17 – Rendement quantique effectif des caméras CPng

5.3.5 Recentrage sous-pixel des photoévénements

Le recentrage sous-pixel d'un photoévènement peut être effectué en temps réel ou en temps différé, grâce à la sauvegarde de la position au pixel près du pic du photoévènement, et des valeurs des pixels dans la boîte 3×3 centrée sur le pic du photoévènement. A cause de la troncature de la fenêtre de centrage et de la lente décroissance spatiale de la forme des photoévénements sur les ailes, le barycentre est un estimateur biaisé vers le centre de la fenêtre de centrage et également sous-optimal pour recentrer les photoévénements, comme on l'a déjà vu à la section 5.1.2.4. Dans le même cadre que la détection par maximum de vraisemblance, une approche de recentrage plus judicieuse consiste à optimiser le critère

$$\chi_{\text{local}}^2(\mathbf{x}) = \sum_{\Delta \mathbf{x} \in S_k} \left(\frac{d(\Delta \mathbf{x}) - \alpha^+(\mathbf{x})s(\Delta \mathbf{x} - \mathbf{x})}{\sigma_{\text{CCD}}} \right)^2, \quad (5.38)$$

où $\alpha^+ = \sum_{\Delta \mathbf{x} \in S} d(\mathbf{x} + \Delta \mathbf{x})s(\Delta \mathbf{x}) / \sum_{\Delta \mathbf{x} \in S} s^2(\Delta \mathbf{x})$. En développant l'équation (5.38), cela conduit à calculer la position recentrée par

$$\mathbf{x}_k^+ = \arg \max_{\mathbf{x}} \frac{\sum_{\Delta \mathbf{x} \in S_k} d(\Delta \mathbf{x}) \times s(\Delta \mathbf{x} - \mathbf{x})}{\sqrt{\sum_{\Delta \mathbf{x} \in S_k} s^2(\Delta \mathbf{x} - \mathbf{x})}}. \quad (5.39)$$

On remarque qu'on retrouve l'expression d'un maximum de corrélation avec la forme du photoévènement, si on néglige au dénominateur de l'expression précédente, les variations de l'énergie du modèle du photoévènement avec la translation \mathbf{x} .

Comme on a affaire à de vraies données, avec des bruits et des erreurs de modélisation éventuellement non négligeables, on peut effectuer un débiaisage a posteriori des positions sous-pixel des photoévénements calculées par maximum de vraisemblance, pour améliorer le centrage. Le raisonnement de correction d'histogramme pour effectuer le débiaisage est effectué à une dimension et appliqué de façon séparable en deux dimensions. Il suppose que les positions sous-pixel pour un grand nombre de photoévénements recentrés sont uniformément réparties à l'intérieur d'un pixel. Soit $f_{\text{exp}}(x_m)$ la densité de probabilité des positions mesurées et $f_{\text{vrai}}(x_r)$ la densité de probabilité des positions réelles, supposée uniforme, de valeur $1/a$ entre $-a/2$ et $+a/2$, où a est la taille du pixel. En définissant la position vraie que l'on souhaite retrouver comme une fonction de la position mesurée :

$$x_r \triangleq x_r(x_m), \quad (5.40)$$

on obtient :

$$\left| \frac{dx_r}{dx_m} \right| = \frac{f_{\text{exp}}(x_m)}{f_{\text{vrai}}(x_r)}, \quad (5.41)$$

De plus comme on souhaite que $x_r(x_m)$ soit une bijection positive de $[-a/2, +a/2]$ sur $[-a/2, +a/2]$, et en utilisant la valeur de $f_{\text{vrai}}(x_r)$ sur $[-a/2, +a/2]$, on obtient :

$$x_r'(x_m) = a \times f_{\text{exp}}(x_m), \quad (5.42)$$

En intégrant cette dernière équation, et puisque $s(-a/2) = -a/2$ et $s(a/2) = a/2$, on trouve que le débiaisage de la position mesurée par maximum de vraisemblance doit être effectué par :

$$x_r(x_m) = a \times \left[F_{\text{exp}}(x_m) - \frac{1}{2} \right], \quad (5.43)$$

où $F_{\text{exp}}(x_m) = \int_{-a/2}^{x_m} f_{\text{exp}}(x'_m) dx'_m$ est la fonction de distribution cumulative des positions mesurées. $F_{\text{exp}}(x_m)$ peut éventuellement être rééchantillonnée à la résolution souhaitée par interpolation de la fonction cumulative mesurée sur les données.

5.3.6 Image des microcanaux

Pour éprouver la technique de recentrage que je viens de présenter, je l'ai appliquée sur un cube d'images, en essayant d'imager les micro-canaux de la galette du deuxième intensificateur. En effet à cause de l'arrangement spatial régulier en hexagones des microcanaux, les photoévénements recentrés doivent recréer cette structure. En théorie la résolution spatiale ultime devrait donc être obtenue en attribuant chaque photoévénement recentré au centre du microcanal le plus proche, et elle est donc limitée par l'espacement entre les centres de deux microcanaux. Cet espacement est de l'ordre de 1/4 de pixel CCD pour CPng, en tenant compte de la focalisation par la lentille. En parvenant à mettre en évidence la structure des microcanaux, je prouve ainsi que la méthode de recentrage permet d'atteindre la résolution spatiale maximale permise par le détecteur, i.e. 1/4 de pixel CCD.

Puisqu'il n'est pas envisageable d'acquérir un volume de données suffisant pour imager directement les microcanaux en longue pose, j'ai cherché à retrouver l'information des microcanaux dans l'autocorrélation moyenne des images. Celle-ci effectue en effet une forme d'intégration spatiale de l'information périodique (ici la structure hexagonale des microcanaux) et nécessite donc un volume plus raisonnable d'images. Comme on le voit sur la figure 5.18, je retrouve dans l'autocorrélation l'information des microcanaux au quart de pixel CCD. Pour estimer le bénéfice de ma méthode par rapport au barycentre et par rapport à la méthode de corrélation incomplète sans débiaisage, j'ai représenté sur la figure 5.19 une même petite partie de l'autocorrélation, obtenue avec les différentes techniques de centrage. Le quadrillage jaune représente la grille des pixels CCD. On voit clairement que le barycentre ne permet pas d'obtenir un recentrage sous-pixel, et on remarque tout aussi nettement que le débiaisage a posteriori permet d'affiner la méthode de recentrage par maximum de vraisemblance, d'après le plus grand nombre de microcanaux et leur meilleur contraste.

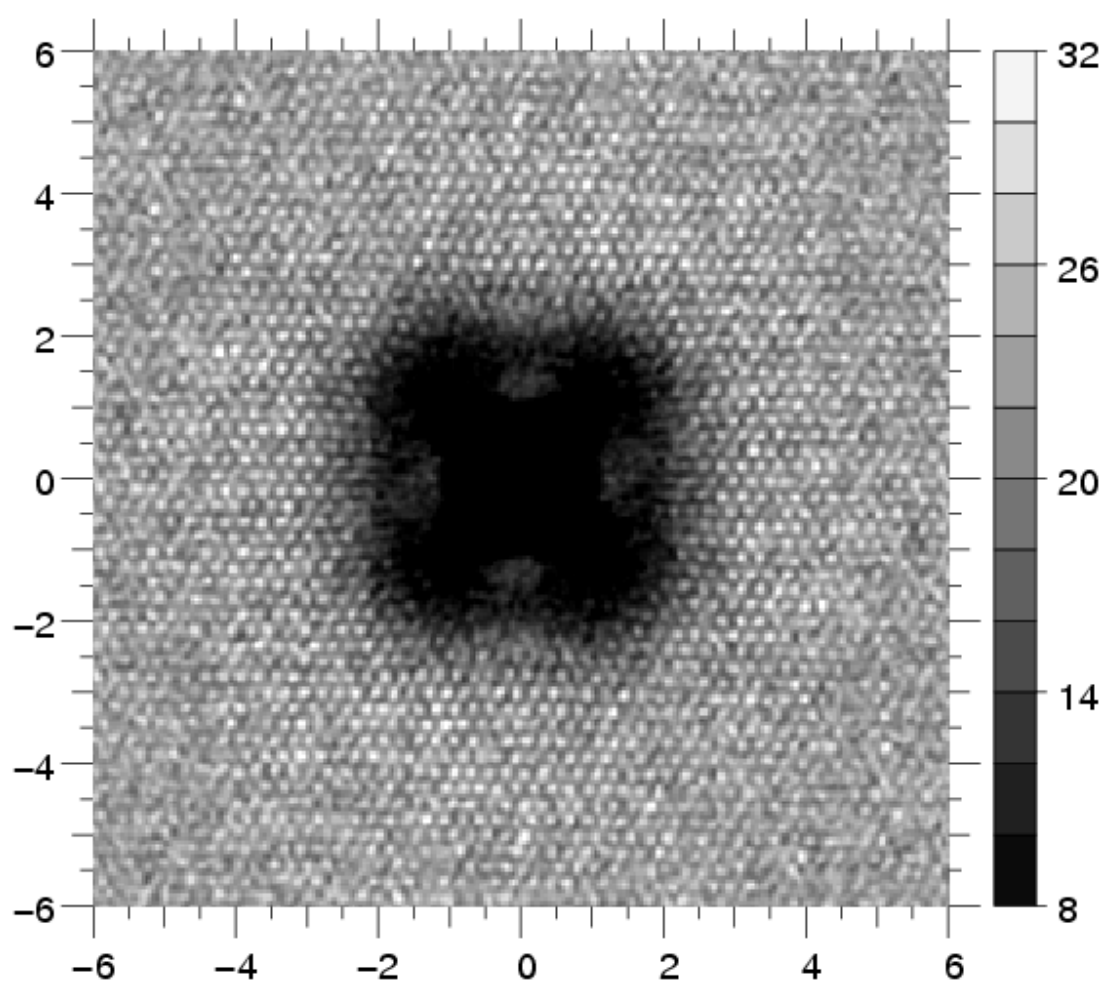


FIG. 5.18 – Autocorrélation moyenne des images en mode comptage et avec les photoévénements recentrés par maximum de vraisemblance+débiaisage a posteriori

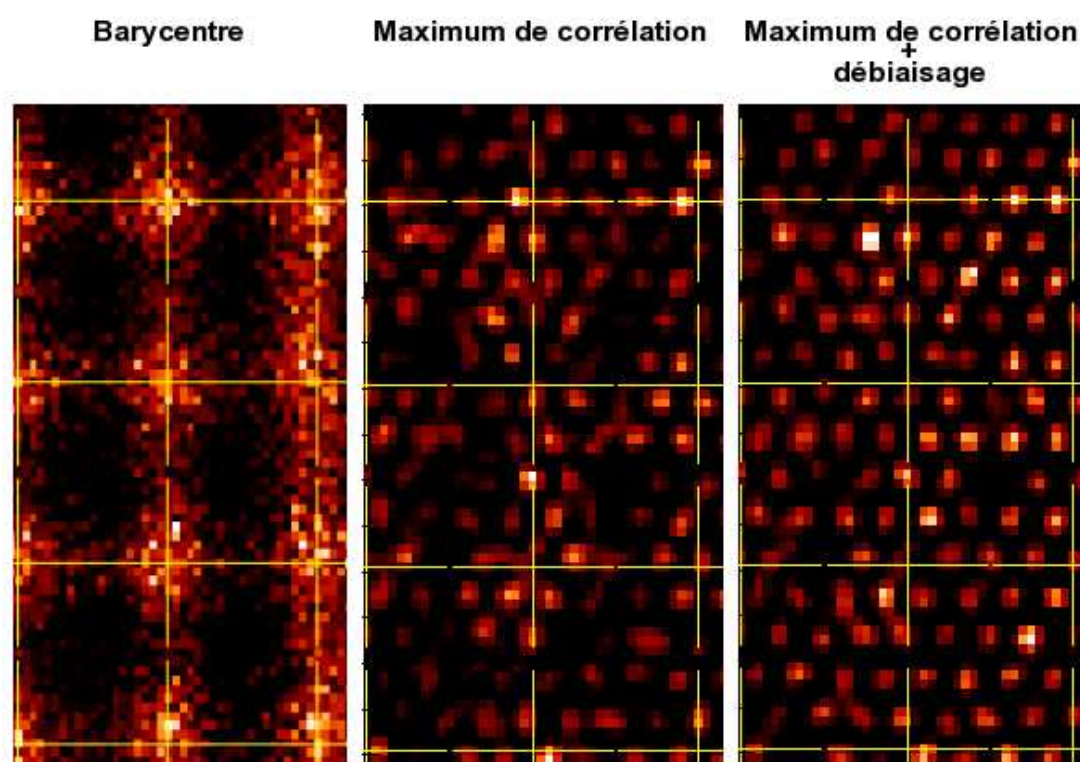


FIG. 5.19 – Zoom sur une partie de l'autocorrélation pour les différentes techniques de centrage

Conclusion

La reconstruction de phase (front d'onde) à partir d'images est un problème inverse non-linéaire difficile avec de nombreuses applications en physique. En particulier en Astronomie, elle peut servir pour estimer les dégradations optiques induites par la turbulence, que ce soit en vue d'un couplage avec une Optique Adaptative ou pour des approches a posteriori de traitement des images de type déconvolution myope ou aveugle. J'ai donc développé au cours de ma thèse un algorithme qui permet de reconstruire la phase de l'onde électromagnétique sur la pupille d'un télescope à partir d'une seule image tavelée au foyer du télescope, grâce à l'optimisation globale du critère bayésien de maximum a posteriori associé à ce problème inverse. L'approche globale que j'ai développée permet d'envisager un nombre beaucoup plus important de degrés de liberté pour un système d'Optique Adaptative, que ce qui a été fait jusqu'à présent avec une seule image, à savoir un passage de 16 à 120 degrés de liberté typiquement, et cela sans même s'appuyer sur une grossière estimation initiale des paramètres de phases que permettraient les corrélations temporelles des fronts d'onde.

J'ai ensuite montré le substantiel bénéfice qu'apporte la diversité polychromatique de phase, c'est-à-dire l'emploi d'images à différentes longueurs d'onde paramétrées par des dégradations communes de phase, pour augmenter encore ce nombre de degrés de liberté. En effet je parviens à reconstruire les fronts d'onde pour des D/r_0 de l'ordre de 70 voire plus grâce à cette technique, ce qui permet, à partir de deux images, de passer de 120 à 5000 (voire plus) degrés de liberté pour un système d'Optique Adaptative.

Enfin dans le cadre du projet d'Etoile Laser Polychromatique, j'ai développé plusieurs méthodes visant à améliorer le centrage de taches images par rapport au barycentre. J'ai montré que la modélisation de la distribution de la tache image turbulente qui est opérée en estimant toutes les aberrations de phase — par la précédente reconstruction de phase par exemple — peut servir à améliorer considérablement la précision de centrage, et permettre un gain de l'ordre du rapport de taille entre la tache image turbulente et la tache de diffraction du télescope. J'ai également développé une méthode de centrage par interspectre bichromatique, qui vise à exploiter un nombre intermédiaire de fréquences spatiales dans le cas où il ne serait pas possible de remonter aux aberrations de phase d'ordres élevés, à cause de l'extension spatiale de l'étoile laser par exemple.

Pour améliorer encore mes algorithmes de reconstruction de phase monochroma-

tique et polychromatique, il me semble intéressant, au vue du bon fonctionnement de méthodes de reconstruction de type POCS, d'examiner le comportement en terme de nombre de minima locaux, d'autres métriques que celle qui découle des fluctuations gaussiennes pour le critère d'adéquation expérimentale. De plus il resterait à comparer les erreurs de reconstruction que j'obtiens, avec une expression analytique de la variance minimale de l'estimateur de maximum a posteriori, qu'on peut obtenir (Fessler [95] par exemple) en supposant que le critère non-linéaire d'adéquation a posteriori est convexe, ce qui est vrai localement, et cela afin de fixer les limites ultimes de précision de la méthode et les comparer à celles que j'obtiens. Enfin pour augmenter encore le nombre de degrés de liberté, il apparaît important de rechercher une meilleur estimation de la pente de la surface d'onde que celle procurée par le barycentre de l'image. A défaut de trouver un bon estimateur, il serait déjà bénéfique d'introduire une légère fluctuation de pente autour du barycentre dans les départs aléatoires.

En ce qui concerne le centrage pour le projet ELP-OA, il reste à comparer plus en détail les performances de ces techniques avec celles d'un centrage effectué dans le plan pupille. En effet comme la précision de centrage plan-pupille est inversement proportionnelle au nombre de sous-pupilles employées, on peut également penser découper la pupille plus grossièrement que $(D/r_0)^2$ sous-pupilles, jusqu'à la limite de résolution de l'étoile laser par les sous-pupilles, et utiliser de la reconstruction de phase éventuellement polychromatique pour centrer les imagerie obtenues. En supposant que la focalisation de l'étoile laser est effectuée à l'aide d'un projecteur laser monolithique concentrant au mieux la tache laser par rapport à la turbulence (i.e. $\sigma_{\text{étoile}} \simeq \lambda/[3 r_0]$), on s'attend à ce que le gain de techniques plan-pupille par rapport au barycentre des images, soit de l'ordre de celui obtenu par l'interspectre bichromatique voire légèrement mieux grâce à la corrélation de l'information entre les différentes voies monochromatiques. Enfin dans le cas où un projecteur laser interférométrique serait utilisé pour conserver les structures les plus fines de la turbulence dans les images, il resterait à combiner sous forme de déconvolution aveugle, mon algorithme de reconstruction de phase polychromatique avec une estimation conjointe de la distribution angulaire d'intensité de l'étoile laser artificielle polychromatique et interférométrique.

Bibliographie

- C. Allen. *Astrophysical Quantities*. University of London, 1976.
- G.R. Ayers, M.J. Northcott, and J.C. Dainty. Knox-thompson and triple-correlation imaging through atmospheric turbulence. *J. Opt. Soc. Am. A*, 5 :963–985, 1988.
- H. W. Babcock. The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of the Pacific*, 65 :229, 1953.
- C. Béchet, M. Tallon, and E. Thiébaud. Frim : minimum-variance reconstructor with a fractal iterative method. In *Proceedings of the SPIE on Advances in Adaptive Optics II*, volume 6272, 2006.
- A. Blanc. *Identification de réponse impulsionnelle et restauration d’images : apports de la diversité de phase*. PhD thesis, Université Paris XI, 2002.
- M. Born and E. Wolf. *Principles of Optics : Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.
- G. Buffon. *Essai d’arithmétique morale*. 1777.
- A. Burian, P. Kuosmanen, J. Saarinen, and C. Rusu. Non-parametric reconstruction of distribution functions from observed galactic discs. *Proceedings of the 2000 IEEE Signal Processing Society Workshop*, 2 :652–661, 2000.
- R.C. Cannon. Global wave-front reconstruction using shack-hartmann sensors. *J. Opt. Soc. Am. A*, 12 :2031–2039, September 1995.
- H. Cramér. *Mathematical Methods of Statistics*. Princeton University Press, 1946.
- W.C. Davidon. Variable metric methods for minimization. Technical report, Argonne National Lab Report, IL, 1959.
- J.E. Dennis and J.J. Moré. A characterization of superlinear convergence and its application to quasi-newton methods. *Math. Comp.*, 28 :549–560, 1974.

- J. Ebersberger and G. Weigelt. Isoplanicity and lifetime of stellar speckle interferograms. *Opt. acta*, 32 :793–801, 1985.
- B Edlén. The refractive index of air. *Metrologia*, 2(2) :71–80, 1966.
- B. Andresen et al. On lumped models for thermodynamic properties of simulated annealing problems. *J. Phys.France*, 49 :1485, 1988.
- J. R. Fienup. Phase retrieval algorithms : a comparison. *Appl. Opt.*, 21(15) :2758–2769, 1982.
- J.R. Fienup. Reconstruction of an object from the modulus of its fourier transform. *Optics Letters*, 3 :27–29, 1978.
- J.R. Fienup and G.B. Feldkamp. Astronomical imaging by processing stellar speckle interferometry data. In *SPIE Proc. : Applications of Speckle Phenomena*, pages 95–102, 1980.
- R. Fletcher and M. Powell. A rapidly convergent descent method for minimization. *Comput. J.*, 6 :163–168, 1963.
- R. Foy. *Optics in Astrophysics*, chapter Laser Guide Stars : Principle, cone effect and tilt measurement, pages 249–271. NATO Science Series, 2005.
- R. Foy. The polychromatic artificial sodium star : a new concept for correcting the atmospheric tilt. *Astronomy and Astrophysics*, 111 :569–578, 1995.
- R. Foy and F. Foy, editors. *Optics in Astrophysics*. NATO Science Series, 2005.
- R. Foy and A. Labeyrie. Feasibility of adaptive telescope with laser probe. *Astron. & Astrophys.*, 152 :29–31, 1985.
- D. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *J. Opt. Soc. Am.*, 56 :1372, 1966.
- D. Fried. Anisoplanatism in adaptive optics. *J. Opt. Soc. Am.*, 72 :52–61, 1982.
- S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6 :671–680, 1984.
- R.W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35(28) :237–246, 1972.
- J.C. Gilbert and J. Nocedal. Global convergence properties of conjugate gradient methods for optimization. Technical report, Rapport de Recherche INRIA-Paris, 1990.

- J. Girard. *Validation sur le ciel du concept d'étoile laser polychromatique*. PhD thesis, Université Claude-Bernard Lyon1, 2005.
- M.S. Gockenbach. *Nonlinear Programming : Theory and Practice*. 2005.
- R.A. Gonsalves. Phase retrieval from modulus data. *J. Opt. Soc. Am.*, 66 :961–964, 1976.
- R.A. Gonsalves. Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21 :829–832, 1982.
- J. Goodman. *Introduction to Fourier optics*. McGraw-Hill Book Companies, 1968.
- A. Griewank and P.L. Toint. chapter On the unconstrained optimization of partially separable objective functions, pages 301–312. M.J.D. Powell eds., Academic press - London, 1981.
- W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57 :97–109, 1970.
- J. Idier, L. Mugnier, and A. Blanc. Statistical behavior of joint least-square estimation in the phase diversity context. *IEEE Transactions on Image Processing*, 14 :2107–2116, 2005. doi : 10.1109/TIP.2005.859365.
- R. Irwan and R.G. Lane. Phase retrieval with prior information. *J. Opt. Soc. Am. A*, 15 :2302–2311, 1998.
- D.P. Karo and A.M. Schneiderman. Speckle interferometry at finite spectral bandwidths and exposure times. *J. Opt. Soc. Am*, 68 :480–485, 1978.
- S.C. Kirkpatrick, D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220 :2784–2796, 1983.
- K.T. Knox and B.J. Thompson. Recovery of images from atmospherically degraded short exposure photographs. *Astrophys. J. Lett.*, 193 :L45–L48, 1974.
- A. Kolmogorov. Local structure of turbulence in incompressible fluids with very high reynolds number. *Dokl. Akad. Nauk. SSSR*, 30 :301–305, 1941.
- D. Korff. Analysis of a method for obtaining near-diffraction-limited information in the presence of atmospheric turbulence. *J. Opt. Soc. Am*, 63 :971–980, 1973.
- A. Labeyrie. Attainment of diffraction-limited resolution in large telescopes by fourier-analyzing speckle patterns in star images. *Astron. Astrophys.*, 6 :85–87, 1970.

- A.-M. Lagrange, Denis Mourard, and Pierre Léna, editors. *High angular resolution in astrophysics*. Kluwer Academic, NATO ASI Series, 1997.
- R. Lane and M. Tallon. Wavefront reconstruction using a shack-hartmann sensor. *Applied Optics*, 31 :6902–6908, 1992.
- R. Lane, R. Clare, and M. van Dam. *Optics in Astrophysics*, chapter chap. 13 : Statistical Wavefront Sensing, pages 375–395. Natos science, 2005.
- R. G. Lane. Methods for maximum-likelihood deconvolution. *J. Opt. Soc. Am. A*, 13 (10) :1992–1998, 1996.
- R.G. Lane. Blind deconvolution of speckle images. *J. Opt. Soc. Am. A*, 9(9) :1508–1514, September 1992.
- R.G. Lane, A. Glindemann, and J. C. Dainty. Simulation of a kolmogorov phase screen. *Wave in random media*, 2 :209–224, 1992.
- F. Martin, R. Conan, and A. Tokovinin et al. Optical parameters relevant for high angular resolution at paranal from gsm instrument and surface layer contribution. *Astron. Astrophys.*, 144 :39–44, 2000.
- N. Metropolis and S.M. Ulam. The monte carlo method. *J. Am. Stat. Assoc.*, 44 : 335–341, 1949.
- N. Metropolis, Rosenbluth, and Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21 :1087–1092, 1953.
- V. Michau, G. Rousset, and J.C. Fontanella. *Real Time and Post Facto Solar Image Correction*, chapter Wavefront Sensing from Extended Sources, pages 124–128. NSO/SP SummerWorkshop Series No. 13, 1992.
- A. Michelson and J. Benoit. Valeur du mètre en longueurs d’ondes. *Trav. Mem. Bur. Int. Poids Mes.*, 11, 1895.
- J.J. Moré and D.C. Sorensen. Computing a trust region step. *SIAM J. Sci. Stat. Comp.*, 4(3) :553–572, 1983.
- K. Mosegaard. *Monte Carlo analysis of inverse problems*. 2006.
- K. Mosegaard and M. Sambridge. Monte carlo analysis of inverse problems. *Inverse Problems*, 18, 2002.
- L. Mugnier, A. Blanc, and J. Idier. *Advances in Imaging & Electron Physics*, chapter chap. 1 : Phase Diversity : a technique for Wave-Front Sensing and for Diffraction-Limited Imaging, pages 1–76. Elsevier, 2006.

- M. Nieto-Vesperinas and J.A. Mendez. Phase retrieval by monte carlo methods. *Opt. commun.*, 59 :249–254, 1986.
- M. Nieto-Vesperinas, R. Navarro, and F.J. Fuentes. Performance of a simulated-annealing algorithm for phase retrieval. *J. Opt. Soc. Am. A*, 5 :30–38, 1988.
- J. Nocedal. Theory of algorithms for unconstrained optimization. *Acta Numerica*, 1992.
- J. Nocedal and J. Wright. *Numerical Optimization*. 2006.
- J.D. Nulton and P. Salamon. Statistical mechanics of combinatorial optimization. *Phys. Rev*, 37 :1351–, 1988.
- A.M. Obukhov. Structure of the temperature field in a turbulent flow. *Izv. Acad. Nauk. SSSR Ser. Geograf. Geofiz.*, 13 :58, 1949.
- A. Papoulis. *Probability, Random Variables and Stochastic Processes*. 1965.
- J.M. Pedersen. *Simulated annealing and finite-time thermodynamics*. PhD thesis, University of Copenhagen, 1990.
- C. Perrier. Amplitude estimation from speckle interferometry. In D. Alloin and J.-M. Mariotti, editors, *NATO ASIC Proc. 274 : Diffraction-Limited Imaging with Very Large Telescopes*, pages 99–+, 1989.
- C. Pichon and E. Thiébaud. Non-parametric reconstruction of distribution functions from observed galactic discs. *Mon. Not. R. Astron. Soc.*, 301 :419, 1998.
- B. Platt and R.V. Shack. Lenticular hartmann screen. *Opt. Sci. Center Newsl.*, 5 : 15–16, 1971.
- E. Polak and G. Ribière. Note sur la convergence de méthodes de directions conjuguées. *Rev. Française Informat Recherche Opérationnelle*, 16 :35–43, 1969.
- R. Ragazzoni. Pupil plane wavefront sensing with an oscillating prism. *Journal of Modern Optics*, 43(2) :289–293, 1996.
- F. Roddier. *The effects of atmospheric turbulence in optical astronomy*, volume 19, pages 281–376. North-Holland Publishing Company, Amsterdam, 1981.
- F. Roddier. Curvature sensing and compensation : a new concept in adaptive optics. *Appl. Opt.*, 27 :1223–1225, 1988.
- F. Roddier, editor. *Adaptive Optics in Astronomy*. Cambridge University Press, 1999.

- F. Roddier, J.M. Gilli, and G. Lund. On the origin of speckle boiling and its effects on stellar speckle interferometry. *Journal of Optics*, 13 :263–271, 1982.
- M. Roggemann, B. Welsh, and R. Fugate. Improving the resolution of ground-based telescopes. *Reviews of Modern Physics*, 69 :437–501, 1997.
- G. Rousset. *Adaptive Optics in Astronomy*, chapter Chapter 5, Wave-front sensors, pages 91–130. Cambridge University Press, 1999.
- S.K. Saha. Modern optical astronomy : technology and impact of interferometry. *Rev. Mod. Phys.*, 74(2) :551–600, 2002.
- R.J. Sasiela. *Electromagnetic Wave Propagation in Turbulence*. 1994.
- J. Skilling and R.K. Bryan. Maximum entropy image reconstruction : general algorithm. *MNRAS*, 211 :111–124, 1984.
- S.M. Stigler. *The history of statistics - The measurement of uncertainty before 1900*. Cambridge, Ma. : Harvard University Press, 1986.
- M. Tallon and R. Foy. Adaptive telescope with laser probe : isoplanatism and cone effect. *Astron. Astrophys.*, 235 :549–557, 1990.
- A. Tarantola. *Inverse Problem Theory - Methods for Data Fitting and Model Parameter Estimation*. Elsevier, 1987.
- V.I. Tatarskii. *Wave Propagation in a Turbulent Medium*. 1961.
- V.I. Tatarskii. *The effects of the turbulent atmosphere on wave propagation*. 1971.
- G.I. Taylor. The spectrum of turbulence. *Proc. R. Soc. Lond. A*, 164 :476–490, 1938.
- J.R. Taylor, B. Alex King III, J. Steincamp, and J. Rakoczy. The spectrum of turbulence. *P.A.S.P.*, 118 :319–323, 2006.
- E. Thiébaud. Speckle interferometry with a photon-counting detector. *Astron. & Astrophys.*, 284 :340–348, 1994.
- P.L. Toint. *Harwell Subroutine Library*, chapter VE08AD, a routine for partially separable optimization with bounded variables. A.E.R.E, U.K., 1983.
- P.L. Toint. Global convergence of the partitioned bfgs algorithm for convex partially separable optimization. *Math. Programming*, 36 :290–306, 1986.
- A. Tokovinin, S. Baumont, and J. Vasquez. Statistics of turbulence profile at cerro tololo. *Mon. Not. R. Astron. Soc.*, 340 :52–58, 2002.

- M. Unser, A. Akram, and M. Eden. Enlargement or reduction of digital images with minimum loss of information. *IEEE trans. image process*, 4 :247–258, 1995.
- J. Vaillant. *Mesure du tilt atmosphérique à partir de sa variation chromatique pour l'étoile laser polychromatique*. PhD thesis, UCBL, 2002.
- J. Vaillant, E. Thiébaud, and M. Tallon. Elpoa : Data processing of chromatic differences of the tilt measured with a polychromatic laser guide star. In P. L. Wiziniowich, editor, *Adaptive Optical Systems Technology*, volume 4007 of *SPIE Conference*, page submitted. Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, 2000.
- M. van Dam and R. Lane. Wavefront slope estimation. *J. Opt. Soc. Am. A*, 17 :1319–1324, 2000.
- M. van Dam and R. Lane. Theoretical performance of phase retrieval on a subdivided aperture. *Optical Engineering*, 41 :1387–1395, 2002a.
- M. van Dam and R.G. Lane. Extended analysis of curvature sensing. *J. Opt. Soc. Am. A*, 19 :745–752, 2002b.
- V.V. Voitsekhovich. Outer scale of turbulence : comparison of different models. *josaa*, 12 :1346–1353, 1995.
- J.G. Walker. Optimum exposure time and filter bandwidth in speckle interferometry. In *Proceedings of the IAU Coll.*, 1978.
- G. Weigelt. Modified astronomical speckle interferometry 'speckle masking'. *Optics Communications*, 21 :55–59, 1977.
- P. Wolfe. Convergence conditions for ascent methods. *SIAM Rev.*, 11 :226–235, 1969.
- P. Wolfe. Convergence conditions for ascent methods ii : Some corrections. *SIAM Rev.*, 13 :185–188, 1971.
- M.C. Won, D. Mnyama, and R.H.T Bates. Improving initial phase estimates for phase retrieval algorithms. *Journal of Modern Optics*, 32 :377–396, 1985.
- A.M. Yaglom. On the local structure of the temperature field in a turbulent flow. *Doklady Acad. Nauk. SSSR Ser. Geograf. Geofiz.*, 69 :73, 1949.

Phase retrieval from speckle images

Xavier Rondeau,* Eric Thiébaud, Michel Tallon, and Renaud Foy

Université de Lyon, F-69000 Lyon, France; Université de Lyon 1, F-69622 Villeurbanne, France; Centre de Recherche Astronomique de Lyon, Observatoire de Lyon, 9 Avenue Charles André, F-69561 Saint-Genis Laval Cedex, France; and CNRS, UMR 5574, Ecole Normale Supérieure de Lyon, F-69007 Lyon, France

*Corresponding author: rondeau@obs.univ-lyon1.fr

Received April 27, 2007; accepted July 9, 2007;
posted July 25, 2007 (Doc. ID 82456); published September 24, 2007

In ground-based astronomy, the inverse problem of phase retrieval from speckle images is a means to calibrate static aberrations for correction by active optics. It can also be used to sense turbulent wavefronts. However, the number of local minima drastically increases with the turbulence strength, mainly because of phase wrapping ambiguities. Multifocal phase diversity has been considered to overcome some ambiguities of the phase retrieval problem. We propose an effective algorithm for phase retrieval from a single focused image. Our algorithm makes use of a global optimization strategy and an automatically tuned smoothness prior to overcome local minima and phase degeneracies. We push the limit of $D/r_0=4$ achieved by Irwan and Lane [J. Opt. Soc. Am. A. **15**, 2302 (1998)] up to $D/r_0=11$, which is a major improvement owing to the drastic increase in the problem complexity. We estimate the performances of our approach from consistent simulations for different turbulence strengths and noise levels (down to 1500 photons per image). We also investigate the benefit of temporal correlation. © 2007 Optical Society of America

OCIS codes: 100.5070, 100.3190, 100.0100, 010.7350.

1. INTRODUCTION

Phase retrieval consists in recovering the phase of an electromagnetic wave from intensity measurements (i.e., images). It is a difficult inverse problem of signal processing that has many applications in physics: electron microscopy, crystallography, wavefront sensing, and blind deconvolution. It has been widely prospected in the case of one or two Fourier intensity measurements by means of error-reduction algorithms and nonlinear least-square (NLLS) fitting algorithms [1–3]. In astronomy, phase retrieval could be used to sense the aberrated turbulent wavefront and thus achieve diffraction-limited resolution for scientific observations. To perform the phase retrieval, a model of the image given the phase in the pupil plane is required, which we detail in Section 2.

In this article, we do not investigate any variant of Fienup's error-reduction algorithms but rather follow the approach of Irwan and Lane [4]: We retrieve the phase model by explicitly maximizing the likelihood of the image data. Since the likelihood is not a quadratic function of the parameters, there is no analytic solution to that problem. The most likely phase estimate is found by successive minimizations of the negative log likelihood f_{ML} , using local quadratic approximations and a NLLS fitting procedure. However some strict phase degeneracies cannot be disentangled with the maximum likelihood (ML) alone. This can be palliated within a Bayesian framework, as justified in Section 3. Indeed, for a symmetrical pupil, the image model is not changed [4] neither when a piston is added to the phase, nor when $2k\pi$ is added to any phase value, nor when $\varphi(\mathbf{r})$ is replaced with $-\varphi(\mathbf{r})$. The two first issues constitute the phase unwrapping

problem. Regularization, as explained below, is expected to do some unwrapping since it enforces the smoothness (hence continuity) of the restored phase. The last point is even worse than solely the ambiguity between $\varphi(\mathbf{r})$ and $-\varphi(\mathbf{r})$, as there is a quasi-exact degeneracy with respect to f_{ML} on the sign of each even mode of a modal phase decomposition. Once again, the regularization is expected to be decisive, because it can disentangle all these degeneracies, except the $\pm\varphi(\pm\mathbf{r})$ one, to which the prior and likelihood penalties are both insensitive.

One of the objectives of this work is to apply phase retrieval under worse turbulent conditions and to larger pupil diameters D than the limit of $D/r_0=4$ reached by Irwan and Lane [4]. Because there are much more local minima with increasing D/r_0 , the success of the local optimization of the Bayesian criterion is uncertain unless a *correct* initial phase is available. The technique of phase diversity, with several intensity measurements in different planes [5], is today widely used to ease phase retrieval problems by avoiding some of these ambiguities. Intended defocus can also help to disentangle some phase degeneracies and allow phase retrieval from a single defocused image [6]. Although there can be many local minima with measurements from a single focused image, we however verified that the global maximum posterior likelihood provides an accurate phase estimate (see Section 6). Solving the phase retrieval problem from a single image, therefore, consists mostly in finding this global optimum. To that end, we derive a dedicated global optimization strategy that is effective in solving the problem in a reasonable computing time (see Section 4).

Our global optimization relies on an efficient local opti-

mization to iteratively improve the phase estimate. At each iteration, we automatically tune the smoothness level so as to avoid phase ambiguities and local minima. We further use a subspace of search directions to cope with the nonlinearity of the problem and the large number of parameters. The resulting local optimization method is detailed in Section 5.

We derive the performances of our phase retrieval algorithm for uncorrelated wavefronts under various noise levels and turbulence conditions. We also consider the case of temporally correlated wavefronts, which would occur in the case of continuous observations. Our results are presented in Section 6.

2. DATA MODEL

A. Image Formation

We first recall the continuous physical model of short-exposure monochromatic images. We denote $\varphi(\mathbf{r})$ as the aberrated wavefront on the pupil of a ground-based telescope at a position \mathbf{r} in the pupil plane. The atmosphere is often considered as an achromatic medium stratified in refractive index \mathcal{N} , and the aberrated wavefront is then written as

$$\varphi(\mathbf{r}) = \frac{2\pi}{\lambda} \delta(\mathbf{r}), \quad (1)$$

where $\delta(\mathbf{r}) = \int_0^\infty \mathcal{N}(\mathbf{r}, z) dz$ is the optical path length through the transfer medium, λ is the wavelength, and z is the position along the line of sight, which is perpendicular to the pupil plane.

Neglecting the scintillation effects of the turbulence, the monochromatic complex amplitude in the pupil for a pointlike source is written as

$$A(\mathbf{r}) = P(\mathbf{r})e^{i\varphi(\mathbf{r})}, \quad (2)$$

where $P(\mathbf{r})$ is the (complex) amplitude transmission of the telescope, the so-called pupil function. Given the complex amplitude in the pupil, the instantaneous monochromatic brightness distribution in the angular direction \mathbf{x} is written as follows [7]:

$$m(\mathbf{x}) = \alpha |a(\mathbf{x})|^2, \quad (3)$$

where $\alpha > 0$ accounts for the monochromatic flux of the source and for a normalization factor and $a(\mathbf{x})$ is, apart from a change of coordinates, the Fourier transform of the complex amplitude in the pupil:

$$a(\mathbf{x}) = \iint A(\lambda \mathbf{u}) e^{+2i\pi \mathbf{x} \cdot \mathbf{u}} d^2 \mathbf{u}, \quad (4)$$

where \mathbf{u} is the spatial frequency. As can be seen from Eqs. (2)–(4), the relation between the speckle image and the aberrated phase is strongly nonlinear. This gives rise to most of the difficulties of the phase retrieval problem, which aims at estimating the phase $\varphi(\mathbf{r})$ by fitting the model $m(\mathbf{x})$ to the observed image $d(\mathbf{x})$.

B. Discretization

Image detection by a CCD involves sampling of the brightness distribution. The discrete model \mathbf{m} of the monochromatic image writes:

$$m_j \stackrel{\text{def}}{=} m(\mathbf{x}_j), \quad (5)$$

where \mathbf{x}_j is the viewing direction of the j th CCD pixel.

Taking $a_j \stackrel{\text{def}}{=} a(\mathbf{x}_j)$, the discretization is continued as follows:

$$m_j = \alpha |a_j|^2 = \alpha \left| \sum_k F_{j,k} A_k \right|^2, \quad (6)$$

where the $F_{j,k}$'s are the complex coefficients of the Fourier transform approximated by a Riemann sum—we further make a circulant approximation and take \mathbf{F} as being the discrete Fourier transform (DFT). The discretized complex amplitude in the pupil plane reads

$$A_k \stackrel{\text{def}}{=} A(\lambda \mathbf{u}_k) = P_k e^{i\varphi_k}, \quad (7)$$

where \mathbf{u}_k is the k th discrete spatial frequency and

$$P_k \stackrel{\text{def}}{=} P(\lambda \mathbf{u}_k), \quad (8)$$

$$\varphi_k \stackrel{\text{def}}{=} \varphi(\lambda \mathbf{u}_k), \quad (9)$$

are, respectively, the sampled pupil function and the sampled monochromatic phase.

Note that the model described by these equations does not account for the spatial filtering by the pixels of the detector and assumes that the source is unresolved. It is sufficient to convolve the model \mathbf{m} by a linear filter to account for a resolved source and for a more realistic detector. The resulting changes in the mathematical derivations presented in this paper would be straightforward and, for sake of simplicity, we do not consider this additional filtering.

C. Phase Parametrization

In the following, we will show that the phase restoration can be restated as a global optimization problem. To help in solving this problem, it may be advantageous to change the phase parametrization and use

$$\varphi_k = \sum_\ell K_{k,\ell} \psi_\ell, \quad (10)$$

or, in matrix notation,

$$\boldsymbol{\varphi} = \mathbf{K} \cdot \boldsymbol{\psi}, \quad (11)$$

where $\boldsymbol{\psi}$ is the vector of the sought parameters and the columns of the matrix \mathbf{K} form a (e.g., modal or zonal) basis of phase functions. Under such a reparametrization, our phase retrieval problem consists in seeking for the best phase parameters $\boldsymbol{\psi}^*$ given the observed image.

In principle, proper choice of \mathbf{K} can be a deciding issue for the solving of the phase retrieval problem. First, the number of sought parameters can be reduced thanks to the reparametrization. Second, since reparametrization influences the condition number of the problem, \mathbf{K} can be seen as a preconditioner. Hence, proper phase reparam-

etrization may speed up the solving of the global optimization problem and reduce the number of ambiguities.

We have tested different function bases to represent the phase (Fourier decomposition, Gram–Schmidt orthonormalized basis of polynomials, and Karhunen–Loève polynomials) and directly the sampled phase φ . According to our numerous tests, these different parametrizations give similar convergence rates. Our final choice was therefore motivated by the minimal number of parameters to account for the features at the smallest scales in the image. Lane and Tallon [8] demonstrated that the Karhunen–Loève polynomials match this requirement. Providing proper sampling is used, a zonal parametrization can be built to match the previous requirement, with the supplementary advantages that a zonal parametrization speeds up the computations (because it yields a sparse matrix \mathbf{K}) and that it simplifies the interpretation of the wrapping ambiguities. For these reasons we finally choose a zonal parametrization with bicubic spline basis functions, and we use subsampling to reduce the number of parameters in the initial stage of our global optimization strategy (see Section 4). In this reparametrization, ψ is the vector of spline coefficients.

D. Noise

A consistent model for real data must at least account for the photon and the CCD noise. After a CCD dark subtraction meant to take off the CCD electronic bias, the data \mathbf{d} are related to the discrete model \mathbf{m} by

$$\mathbf{d} = \mathbf{m} + \mathbf{n}, \quad (12)$$

where \mathbf{n} accounts for modeling errors and noise contributions: photon noise of the signal (which follows Poissonian statistics) plus detector noise. The detector noise consists in quantization errors by the analog-to-digital converters plus CCD dark current and readout noise (approximately with Gaussian statistics).

3. BAYESIAN FRAMEWORK

The parameters $\{\alpha, \psi\}$ are the unknowns of our problem. In this section, we derive a proper criterion to qualify the best set of parameters given an observed image \mathbf{d} and some prior information.

A. Bayesian Approach

Taking into consideration the statistics of the errors \mathbf{n} , the ML parameters $\{\alpha_{\text{ML}}, \psi_{\text{ML}}\}$ are the ones that maximize the probability of the data given the model

$$\{\alpha_{\text{ML}}, \psi_{\text{ML}}\} = \arg \max_{\alpha, \psi} \Pr(\mathbf{d} | \mathbf{m}(\alpha, \psi)) \quad (13)$$

$$= \arg \min_{\alpha, \psi} f_{\text{ML}}(\alpha, \psi), \quad (14)$$

where the model \mathbf{m} is a function of the parameters $\{\alpha, \psi\}$ and f_{ML} is the ML penalty:

$$f_{\text{ML}}(\alpha, \psi) = -\log \Pr(\mathbf{d} | \mathbf{m}(\alpha, \psi)), \quad (15)$$

up to an additive constant. This approach, however, neither smoothes out the noise [9,10], nor is robust with re-

spect to the ill-conditioning of the inverse problem, nor helps to avoid degeneracies.

Staying in a Bayesian framework, a better approach is to seek for the model parameters that maximize the probability of the model given the data. The resulting so-called maximum *a posteriori* (MAP) solution is written as

$$\{\alpha_{\text{MAP}}, \psi_{\text{MAP}}\} = \arg \max_{\alpha, \psi} \Pr(\mathbf{d} | \mathbf{m}(\alpha, \psi)) \Pr(\mathbf{m}(\alpha, \psi)) \quad (16)$$

$$= \arg \min_{\alpha, \psi} [f_{\text{ML}}(\alpha, \psi) + f_{\text{prior}}(\psi)], \quad (17)$$

where f_{ML} is given by Eq. (15) and the negative logarithm of the prior probability of the phase parameters is

$$f_{\text{prior}}(\psi) = -\log[\Pr(\mathbf{q}(\psi))], \quad (18)$$

where \mathbf{q} is the unscaled model image such that $\mathbf{m} = \alpha \mathbf{q}$; that is, $q_j = |a_j|^2$ [see Eq. (6)]. Note that there is no need to further constrain the flux parameter α . In inverse problem jargon, f_{prior} is also termed the regularization.

In practice, the prior statistic is not exactly known and it is customary to tune the level of regularization by minimizing

$$f(\psi) = f_{\text{ML}}(\psi) + \mu f_{\text{prior}}(\psi), \quad (19)$$

where $\mu > 0$ is a hyper-parameter used to set the relative strength of the *a priori* constraints with respect to the data constraints. Alternatively, μ can be seen as a Lagrange multiplier tuned to match a constraint about the expected value of f_{ML} or f_{prior} . In our global optimization strategy (see Section 4), we make use of an another advantage of the hyper-parameter μ : It can be used to overregularize the solution so as to speed up the initial convergence of the optimization algorithm and to smooth out some of the degeneracies of the problem.

B. Likelihood Term

When data are corrupted by Poissonian and Gaussian noise, assuming Gaussian statistics for the noise is a robust approach [11]. Under this approximation, the ML penalty term reads

$$f_{\text{ML}}(\alpha, \psi) = \frac{1}{2} (\mathbf{d} - \mathbf{m})^T \cdot \mathbf{W}_{\text{ML}} \cdot (\mathbf{d} - \mathbf{m}), \quad (20)$$

where $\mathbf{W}_{\text{ML}} = \mathbf{C}_{\mathbf{n}}^{-1}$ is the inverse of the covariance matrix $\mathbf{C}_{\mathbf{n}} = \langle \mathbf{n} \cdot \mathbf{n}^T \rangle$ of the errors.

As there is no prior constraints on the normalization factor α and since the model is linear with respect to α , it is possible to get rid of this unknown to have a modified criterion that depends only on the phase parameters ψ :

$$f_{\text{ML}}(\psi) = \frac{1}{2} [\mathbf{d} - \alpha^+(\psi) \mathbf{q}(\psi)]^T \cdot \mathbf{W}_{\text{ML}} \cdot [\mathbf{d} - \alpha^+(\psi) \mathbf{q}(\psi)],$$

where \mathbf{q} is the unscaled model image as in Eq. (18) and where

$$\alpha^+(\psi) = \frac{\mathbf{q}(\psi)^T \cdot \mathbf{W}_{\text{ML}} \cdot \mathbf{d}}{\mathbf{q}(\psi)^T \cdot \mathbf{W}_{\text{ML}} \cdot \mathbf{q}(\psi)} \quad (21)$$

is the value of α that minimizes $f_{\text{ML}}(\alpha, \psi)$ given ψ .

Owing to the high nonlinearity of the model with respect to the phase parameters, the unknown ψ cannot be obtained directly and iterative minimization of the penalty is needed. As discussed later, iterative optimization methods require computation of the partial derivatives of the criterion with respect to the parameters. Since

$f_{\text{ML}}(\psi) \stackrel{\text{def}}{=} f_{\text{ML}}(\alpha^+(\psi), \psi)$, by applying the chain rule, we have

$$\frac{\partial f_{\text{ML}}(\psi)}{\partial \psi_k} = \left. \frac{\partial f_{\text{ML}}(\alpha, \psi)}{\partial \psi_k} \right|_{\alpha=\alpha^+(\psi)} + \frac{\partial \alpha^+(\psi)}{\partial \psi_k} \left. \frac{\partial f_{\text{ML}}(\alpha, \psi)}{\partial \alpha} \right|_{\alpha=\alpha^+(\psi)}.$$

By definition, $\alpha^+(\psi)$ minimizes $f_{\text{ML}}(\alpha, \psi)$; hence

$$\frac{\partial f_{\text{ML}}(\alpha, \psi)}{\partial \alpha} = 0 \quad \text{for} \quad \alpha = \alpha^+(\psi).$$

The partial derivatives with respect to the phase parameters then simplify to

$$\frac{\partial f_{\text{ML}}(\psi)}{\partial \psi_k} = \left. \frac{\partial f_{\text{ML}}(\alpha, \psi)}{\partial \psi_k} \right|_{\alpha=\alpha^+(\psi)}, \quad (22)$$

which requires only to compute the partial derivatives of $f_{\text{ML}}(\alpha, \psi)$ with respect to ψ_k and to replace α with its optimal value α^+ given by Eq. (21).

C. Regularization

Since the turbulent phase (after removal of the piston) follows Gaussian statistics with zero mean [7,12], the proper definition of the regularization term in the Bayesian framework is

$$f_{\text{prior}}(\varphi) = \frac{1}{2} \varphi^T \cdot \mathbf{C}_\varphi^{-1} \cdot \varphi, \quad (23)$$

where $\mathbf{C}_\varphi = \langle \varphi \cdot \varphi^T \rangle$ is the covariance matrix of the phase φ . If phase reparametrization is used with $\varphi = \mathbf{K} \cdot \psi$, then

$$f_{\text{prior}}(\psi) = \frac{1}{2} \psi^T \cdot \mathbf{K}^T \cdot \mathbf{C}_\varphi^{-1} \cdot \mathbf{K} \cdot \psi. \quad (24)$$

Note that $\mathbf{K}^T \cdot \mathbf{C}_\varphi^{-1} \cdot \mathbf{K} = \mathbf{C}_\psi^{-1}$ is the inverse covariance matrix of the phase parameters ψ .

As shown by Wallner [13], there is a simple expression that links the covariance matrix \mathbf{C}_φ of a piston-removed turbulent phase φ to the Kolmogorov structure function

$$D_\varphi(\mathbf{r}_i, \mathbf{r}_j) \simeq 6.88 \times \left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{r_0} \right)^{5/3}, \quad (25)$$

where r_0 is Fried's parameter. As the tip-tilt in the images may not be due only to the turbulence, but also, for example, to the instrumentation, we instead used the regularization for the phase with piston and tip-tilt removed.

4. GLOBAL OPTIMIZATION STRATEGY

Local optimization of the MAP criterion $f = f_{\text{ML}} + \mu f_{\text{prior}}$ may be able to avoid some small local minima and ambiguities thanks to the smoothness enforced by the regular-

ization, but it is unable to ensure the global optimization of f because of the highly nonlinear nature of the penalty f_{ML} . Expecting that the global minimum yields a good phase estimate for a given data set, a possible strategy would be to attempt different initial phase starts followed by local optimization of the penalty toward a close local minimum. Among all the local minima found, the deepest one corresponds to the best solution so far. Owing to the large number of parameters that describe the phase about $(D/r_0)^2$, this strategy may, however, take a prohibitively long time to effectively reach the global optimum.

Fortunately, we have been able to make use of the following properties (observed for simulated data) to obtain a good phase estimate in a much shorter computation time:

- We have verified that, in practice, the value of f_{ML} , which measures the agreement of the model with the data, is also a good indicator of the quality of the recovered phase in the sense that if f_{ML} is small enough, then the corresponding phase may not be the best one but is acceptably close to the true phase to the extent of some modulo- 2π wrapping. This is not unexpected once it is realized that there are much fewer parameters to describe the phase about $(D/r_0)^2$ than there are significant measurements [i.e., with sufficient signal-to-noise ratio (SNR)]. Hence a good agreement of the model with the data is a very selective criterion.

- Our nonlinear minimization algorithm described in Section 5 performs only local optimization but is quite efficient at this task: After a limited number of iterations (and well before convergence), it is possible to guess whether or not the algorithm will finally converge to a good minimum by comparing the current value of f_{ML} to some goodness-of-fit threshold $f_{\text{ML}}^{\text{min}}$.

- Restricting the phase sampling to approximately $(D/r_0)^2$ parameters enforces its smoothness. Overregularizing has similar consequences. We can make use of these properties to avoid most of the wrapping ambiguities and to smooth out some of the local minima.

Hence our global optimization strategy is as follows:

1. We choose a threshold value $f_{\text{ML}}^{\text{min}}$; a typical choice is to take $f_{\text{ML}}^{\text{min}} \simeq 10 \times \langle f_{\text{ML}} \rangle$, where $\langle f_{\text{ML}} \rangle = \frac{1}{2} N_{\text{data}}$ is the expected value of the likelihood term.

2. We select a random initial phase described by about $(D/r_0)^2$ phase parameters and perform local optimization. The tip-tilt in the image model is also bounded during the very first iterations, thanks to what can best be done from centroiding, in order to favor the fit of the speckle pattern of the image by higher phase modes.

3. If $f_{\text{ML}} \leq f_{\text{ML}}^{\text{min}}$ after a small number of iterations (typically 50), we continue until convergence, with full phase sampling, which typically requires about 100 iterations. Otherwise, we restart the search with a new random initial phase and proceed with step 2.

As shown by our results in Section 6, this strategy is effective for getting accurate phase estimates. Other specificities of the problem can sometimes be used to speed up our global optimization strategy. In the case of time-correlated wavefronts, we tested that if we take the

previously retrieved phase as an initial phase guess, provided the image acquisition is fast enough, phase retrieval will require only our local optimization.

5. LOCAL MINIMIZATION

A. Newton Method

Because direct minimization of $f(\psi)$ is not possible, local optimization iteratively improves the set of parameters ψ by taking adequate parameter steps $\delta\psi$. Starting from an initial estimate $\psi^{(0)}$, the estimated parameters at the n th iteration read

$$\psi^{(n)} = \psi^{(n-1)} + \delta\psi^{(n-1)}. \quad (26)$$

When the penalty function $f(\psi)$ is known analytically, its partial derivatives can be computed and efficient numerical methods derived from the Newton method [14] can be applied to estimate a good step $\delta\psi$. The Newton method is based on second-order Taylor's expansion of the penalty variation for a small change of parameters:

$$\begin{aligned} \overset{\text{def}}{\delta f(\psi, \delta\psi)} &= f(\psi + \delta\psi) - f(\psi) \\ &= \mathbf{g}(\psi)^T \cdot \delta\psi + \frac{1}{2} \delta\psi^T \cdot \mathbf{H}(\psi) \delta\psi + o(\|\delta\psi\|^2), \end{aligned} \quad (27)$$

where $\mathbf{g}(\psi) = \nabla f(\psi)$ and $\mathbf{H}(\psi) = \nabla^2 f(\psi)$ are, respectively, the gradient (vector of partial derivatives) and Hessian (matrix of partial second derivatives) of $f(\psi)$. Then the quadratic approximation

$$\overset{\text{def}}{\delta f_{\text{quad}}(\psi, \delta\psi)} = \mathbf{g}(\psi)^T \cdot \delta\psi + \frac{1}{2} \delta\psi^T \cdot \mathbf{H}(\psi) \cdot \delta\psi \quad (28)$$

of the local behavior of $\delta f(\psi, \delta\psi)$ can be used to derive an optimal parameter step:

$$\begin{aligned} \delta\psi_{\text{quad}}(\psi) &= \arg \min_{\delta\psi} \delta f_{\text{quad}}(\psi, \delta\psi) \\ &= -\mathbf{H}(\psi)^{-1} \cdot \mathbf{g}(\psi), \end{aligned} \quad (29)$$

providing $\delta\psi_{\text{quad}}$ is small enough to neglect the nonquadratic term $o(\|\delta\psi_{\text{quad}}\|^2)$ in Taylor's expansion (27).

To ensure that the step length $\|\delta\psi\|$ remains sufficiently small, trust-region methods [14] define the optimal step as

$$\delta\psi_{\text{TR}}(\psi) = \arg \min_{\delta\psi} \delta f_{\text{quad}}(\psi, \delta\psi) \quad \text{s.t.} \quad \|\delta\psi\| \leq \Delta, \quad (30)$$

where Δ is the radius of the so-called trust-region, inside which the quadratic approximation is consistent. In practice, the trust-region step can be written as [15]

$$\delta\psi_{\text{TR}}(\psi) = \arg \min_{\delta\psi} \left\{ \delta f_{\text{quad}}(\psi, \delta\psi) + \frac{\lambda}{2} \delta\psi^T \cdot \mathbf{Q} \cdot \delta\psi \right\}, \quad (31)$$

$$= -[\mathbf{H}(\psi) + \lambda \mathbf{Q}]^{-1} \cdot \mathbf{g}(\psi), \quad (32)$$

where \mathbf{Q} is a positive definite matrix that implements the Euclidean norm, i.e., $\|\delta\psi\|^2 = \delta\psi^T \cdot \mathbf{Q} \cdot \delta\psi$, and where λ is a Lagrange multiplier tuned so that the constraint $\|\delta\psi_{\text{TR}}(\psi)\| \leq \Delta$ can be fulfilled. We use the algorithm of Moré and Sorensen [15] to derive the value of the Lagrange multiplier λ .

The main drawback of the (trust-region) Newton method is that it requires computation and inversion of the Hessian matrix, which yields two problems: (i) the second derivatives of $f(\psi)$ must be estimated; and (ii) the $N \times N$ Hessian matrix, where N is the number of sought parameters, must be stored and inverted, which becomes impracticable for a large number of parameters. Since the penalty functions f_{ML} and f_{prior} (and hence f) are the sum of squares, a NLLS version of the Newton method can be used to approximate the Hessian matrix that requires only the computation of the Jacobian matrix (see Appendix A). In the next section, we describe how to solve the second problem by computing the iteration step in a local subspace made of a small number of search directions.

B. Optimization in a Local Subspace

The Newton method for NNLS (see Subsection 5.A and Appendix A), can be applied only for a moderate number of degrees of freedom, that is, for small D/r_0 ratios. Indeed, it requires calculating and storing an $N \times N$ approximation of the Hessian matrix, where the number of parameters N scales as $(D/r_0)^2$. Computation of this approximation makes use of the Jacobian matrix of partial derivatives of the model with respect to the parameters, which has $M \times N$ elements, where M is the number of pixels in the image. Hence the size of the matrices involved in the NLLS Newton method scales as $\max(M, N) \times N$.

When either M and/or N becomes too large to store or to compute such matrices, limited memory variants of the Newton method must be used. For instance, conjugate gradients or a limited memory variable metric method such as L-BFGS [16,17] are algorithms that can be used to solve large nonlinear optimization problems. The phase retrieval problem is, however, so highly nonlinear that the convergence rate of these algorithms will be hardly better than linear. In such cases, Skilling and Bryan [18] have proposed to solve the optimization problem over a linear space made of a small number of search directions. By properly choosing these search directions, it is possible to maintain a fast rate of convergence even with very few directions (three to five are sufficient).

In the local subspace approach, the parameter step is rewritten as

$$\delta\psi = \mathbf{S} \cdot \boldsymbol{\beta} = \sum_i \beta_i \mathbf{s}_i, \quad (33)$$

for a limited number of search directions \mathbf{s}_i , which are the columns of the matrix \mathbf{S} . Then the quadratic approximation of the variation of the penalty function becomes

$$\delta f_{\text{quad}}(\psi, \boldsymbol{\beta}, \mu) = \frac{1}{2} \boldsymbol{\beta}^T \cdot \mathbf{A}(\psi, \mu) \cdot \boldsymbol{\beta} - \mathbf{b}(\psi, \mu)^T \cdot \boldsymbol{\beta}, \quad (34)$$

with

$$\mathbf{A}(\psi, \mu) = \mathbf{S}^T \cdot \mathbf{H}_{\text{ML}}(\psi) \cdot \mathbf{S} + \mu \mathbf{S}^T \cdot \mathbf{H}_{\text{prior}}(\psi) \cdot \mathbf{S}, \quad (35)$$

$$\mathbf{b}(\psi, \mu) = -\mathbf{S}^T \cdot \mathbf{g}_{\text{ML}}(\psi) - \mu \mathbf{S}^T \cdot \mathbf{g}_{\text{prior}}(\psi). \quad (36)$$

A fast way of computing the coefficients of the matrix $\mathbf{A}(\psi, \mu)$ and of the vector $\mathbf{b}(\psi, \mu)$ is detailed in Appendix B. Since Eq. (34) is quadratic with respect to $\boldsymbol{\beta}$, the trust-region Newton method can be used to derive a good set of parameters $\boldsymbol{\beta}_{\text{TR}}$. However \mathbf{Q} must be replaced by $\mathbf{S}^T \cdot \mathbf{Q} \cdot \mathbf{S}$

in order to implement the same metric. Then $\mathbf{S} \cdot \boldsymbol{\beta}_{\text{TR}}$ is the step given by the local subspace approximation of the Newton method:

$$\delta\boldsymbol{\psi}(\mu, \lambda) = -\mathbf{S} \cdot [\mathbf{A}_{\text{ML}} + \mu\mathbf{A}_{\text{prior}} + \lambda\mathbf{S}^T \cdot \mathbf{Q} \cdot \mathbf{S}]^{-1} \cdot [\mathbf{b}_{\text{ML}} + \mu\mathbf{b}_{\text{prior}}], \quad (37)$$

with $\mathbf{A}_{\text{ML}} = \mathbf{S}^T \cdot \mathbf{H}_{\text{ML}} \cdot \mathbf{S}$, $\mathbf{A}_{\text{prior}} = \mathbf{S}^T \cdot \mathbf{H}_{\text{prior}} \cdot \mathbf{S}$, $\mathbf{b}_{\text{ML}} = -\mathbf{S}^T \cdot \mathbf{g}_{\text{ML}}$, and $\mathbf{b}_{\text{prior}} = -\mathbf{S}^T \cdot \mathbf{g}_{\text{prior}}$.

We now explain how we choose the search directions. For instance, making use of the Newton step in Eq. (29) and assuming that $\mathbf{B} \sim \mathbf{H}(\boldsymbol{\psi})^{-1}$ is an approximation of the inverse Hessian that can be carried out in a manageable number of operations, a search direction of choice is the quasi-Newton step:

$$\begin{aligned} \delta\boldsymbol{\psi}_{\text{QN}} &= -\mathbf{B} \cdot \mathbf{g}(\boldsymbol{\psi}) \\ &= -\mathbf{B} \cdot \mathbf{g}_{\text{ML}}(\boldsymbol{\psi}) - \mu\mathbf{B} \cdot \mathbf{g}_{\text{prior}}(\boldsymbol{\psi}). \end{aligned} \quad (38)$$

Following Skilling and Bryan [18] prescription, we take

$$\mathbf{B} = \frac{1}{\mu} \mathbf{H}_{\text{prior}}^{-1}, \quad (39)$$

which depends only on μ . With this choice, the quasi-Newton step reads

$$\delta\boldsymbol{\psi}_{\text{QN}} = -\frac{1}{\mu} \mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{g}_{\text{ML}}(\boldsymbol{\psi}) - \mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{g}_{\text{prior}}(\boldsymbol{\psi}). \quad (40)$$

Since, as explained in Subsection 5.C, μ must be tuned at every iteration, $\delta\boldsymbol{\psi}_{\text{QN}}$ is a linear combination of

$$\mathbf{s}_1 = -\mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{g}_{\text{ML}}(\boldsymbol{\psi}), \quad (41)$$

$$\mathbf{s}_2 = -\mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{g}_{\text{prior}}(\boldsymbol{\psi}), \quad (42)$$

which readily provide the two first search directions for the expansion in Eq. (33). Owing to the highly nonlinear behavior of the penalty, the \mathbf{s}_i 's must be derived at every iteration; however, note that $\mathbf{H}_{\text{prior}}^{-1}$ does not depend on $\boldsymbol{\psi}$ and must be calculated only once.

Additional search directions can be derived as follows: If a step $\delta\boldsymbol{\psi}$ is taken, according to Taylor's expansion, the gradient becomes

$$\mathbf{g}(\boldsymbol{\psi} + \delta\boldsymbol{\psi}) \simeq \mathbf{g}(\boldsymbol{\psi}) + \mathbf{H}(\boldsymbol{\psi}) \cdot \delta\boldsymbol{\psi}, \quad (43)$$

and the *next* quasi-Newton step would therefore be written as

$$\delta\boldsymbol{\psi}_{\text{QN}}^{\text{next}} = -\mathbf{B} \cdot \mathbf{g}(\boldsymbol{\psi} + \delta\boldsymbol{\psi}_{\text{QN}}) \quad (44)$$

$$\simeq \delta\boldsymbol{\psi}_{\text{QN}} - \mathbf{B} \cdot \mathbf{H}(\boldsymbol{\psi}) \cdot \delta\boldsymbol{\psi}_{\text{QN}}, \quad (45)$$

which induces that the secondary direction to consider is $-\mathbf{B} \cdot \mathbf{H}(\boldsymbol{\psi}) \cdot \delta\boldsymbol{\psi}_{\text{QN}}$. Given a search direction of \mathbf{s}_i , these considerations lead to append $-\mathbf{B} \cdot \mathbf{H}_{\text{ML}} \cdot \mathbf{s}_i$ and $-\mathbf{B} \cdot \mathbf{H}_{\text{prior}} \cdot \mathbf{s}_i$ to the list of search directions. However, since $\mathbf{B} \propto \mathbf{H}_{\text{prior}}^{-1}$, the latter does not provide a new search direction and finally we take

$$\mathbf{s}_3 = -\mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{H}_{\text{ML}} \cdot \mathbf{s}_1, \quad (46)$$

$$\mathbf{s}_4 = -\mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{H}_{\text{ML}} \cdot \mathbf{s}_3. \quad (47)$$

It is worth noting that \mathbf{s}_3 and \mathbf{s}_4 depend on the Hessian of f_{ML} , but neither require it to be explicitly calculated nor stored. Indeed \mathbf{H}_{ML} needs to be applied only to a vector, which can be done in an efficient way as detailed in Appendix B. Finally, to prevent the step $\delta\boldsymbol{\psi}^{(n)}$ derived at the n th iteration from canceling any improvements made by the previous step $\delta\boldsymbol{\psi}^{(n-1)}$, we explicitly use the previous step as the last search direction [19]:

$$\mathbf{s}_5 = \delta\boldsymbol{\psi}^{(n-1)}. \quad (48)$$

Given the search directions, we explain in the next section how to tune the value of the hyper-parameter μ .

C. Tuning the Regularization Weight

Our derivation of the criterion $f = f_{\text{ML}} + \mu f_{\text{prior}}$ is based on Bayesian considerations. In this framework, there should be no reason to tune the relative weight of the prior. This would, however, require that there be good estimates of the statistics of the noise and of the turbulence. In practice, precise measurements of Fried's parameter r_0 , of the CCD readout noise, of the CCD gain, etc., may not be available, and a proper mean to adjust the relative regularization weight is needed. Besides, some phase ambiguities, such as modulo- 2π wrapping, can be avoided by imposing enough smoothness for the phase. In our framework, this can easily be achieved by using a high enough value for the hyper-parameter μ in Eq. (23) because the Kolmogorov prior enforces the smoothness of the phase.

For these reasons, during the iterations of our local minimization algorithm, we control the value of the regularization level μ so that it is low enough to yield a reduction of the likelihood term but still high enough to ensure that the choice of the parameter step $\delta\boldsymbol{\psi}$ is noticeably driven by the regularization. To that end, we tune the value of μ so that the expected reduction of the likelihood term is a fraction $\varepsilon \in [0, 1]$ of the best expected reduction that is achievable when regularization is turned off:

$$\mu \quad \text{s.t.} \quad \delta\mathcal{J}_{\text{ML}}^{\text{quad}}[\boldsymbol{\psi}, \delta\boldsymbol{\psi}(\mu)] = \varepsilon \delta\mathcal{J}_{\text{ML}}^{\text{quad}}[\boldsymbol{\psi}, \delta\boldsymbol{\psi}(\mu=0)], \quad (49)$$

where $\delta\mathcal{J}_{\text{ML}}^{\text{quad}}(\boldsymbol{\psi}, \delta\boldsymbol{\psi})$ is the quadratic approximation of the variation of the likelihood term and $\delta\boldsymbol{\psi}(\mu) = \delta\boldsymbol{\psi}(\mu, \lambda=0)$ [see Eq. (37)]. This kind of strategy has proved effective for other nonlinear inverse problems [18,19]. In practice, we take $\varepsilon \approx 0.5$. Note that the trust-region Lagrange multiplier λ in Eq. (37) is tuned after having derived the regularization level μ . To solve the problem in Eq. (49) for μ , we use the bisection algorithm with automatic updating of the bounds implemented by Pichon and Thiébaud [19].

6. RESULTS

In this section we demonstrate the ability of our global optimization algorithm in retrieving wavefronts for various turbulence strengths and noise levels. For each considered D/r_0 ratio and mean photon count, we run our algorithm on at least 250 different simulated wavefronts and images to derive meaningful statistics. The quality of the recovered phase is given by the rms error with respect to

the true phase. We first present results for uncorrelated wavefronts, which is the most difficult case since no information other than a single short-exposure image is available to retrieve the phase. We then show that thanks to the temporal correlation of real turbulence and providing the time between consecutive images is short enough, the previously recovered phase is a *correct* initial estimate (only local minimization is needed).

A. Simulated Data

The Kolmogorov aberrated wavefronts are generated for a 33×33 degree-of-freedom pupil, which enables us to simulate D/r_0 up to about 11, and we account for a central obscuration of $1/4$ th by a secondary mirror, in the design of the simulated pupil. The wavefronts are simulated by a method [20] similar to the midpoint algorithm of Lane *et al.* [21]. We take 75×75 pixels in the images to correctly sample the speckles (at least 2 pixels per speckle).

We use a Poissonian pseudorandom generator for the photon noise and a Gaussian one to simulate the additive CCD readout and quantization noises. Since we want to test the algorithm under low light levels, we assume that the detector has a low readout noise of $\sigma_{\text{CCD}} = 0.32$ photons/pixel comparable to that of an electron-multiplying CCD (EMCCD). However, we note that using a real EMCCD would result in doubling the variance of the signal noise. In our simulations, the photon noise is the prevailing source of noise.

The flux is diluted over more or fewer speckles in the images with respect to the turbulence strength. Hence for a given average number of photons N_{ph} per frame, the relative noise level for the phase retrieval, e.g., the average noise for a speckle, will also depend on the turbulence strength D/r_0 . That is why we introduce the following SNR per speckle and compute it for our different tests:

$$\text{SNR} = \frac{N_{\text{ph}}/N_{\text{sp}}}{\sqrt{N_{\text{ph}}/N_{\text{sp}} + \sigma_{\text{CCD}}^2 \times s_{\text{sp}}}}, \quad (50)$$

where $N_{\text{sp}} \propto (D/r_0)^2$ is the number of speckles and $s_{\text{sp}} \propto (\lambda/D)^2$ is the mean number of pixels per speckle.

B. Uncorrelated Wavefronts

We first test the ability of our phase retrieval algorithm to retrieve uncorrelated wavefronts. This is the worst possible case since, for each considered image, we cannot rely on an initial *correct* phase guess. Global optimization is required, which we achieve thanks to the random phase starts strategy described in Section 4.

Since the algorithm cannot distinguish between the two phases $\varphi(\mathbf{r})$ and $-\varphi(-\mathbf{r})$ because their penalties (likelihood plus regularization) are exactly the same for a symmetrical pupil, we compute the phase retrieval error by choosing between $\varphi(\mathbf{r})$ and $-\varphi(-\mathbf{r})$ in terms of which yields the least error with respect to the true phase. In practice, this is not a problem since very small temporal continuity between the wavefronts favors the choice of the correct one.

We have performed three main sets of simulation for uncorrelated wavefronts. We have first attempted phase retrieval for images with the same mean flux of N_{ph}

Table 1. Parameters Used in Our Different Simulations

Id.	N_{ph}	D/r_0	SNR	Id.	N_{ph}	D/r_0	SNR
1	50,000	5	44.7	7	50,000	8	27.9
2	50,000	7	31.9	8	22,000	8	18.5
3	50,000	9	24.8	9	9000	8	11.8
5	50,000	10	22.4	10	4000	8	7.9
6	50,000	11	18.1	11	1500	8	4.8

^aAt least 250 different simulations were computed for each considered case. Id. denotes identification number. The column N_{ph} gives the mean flux in photons per image, the turbulence strength is quantified by column D/r_0 , and the SNR is the signal-to-noise ratio per speckle as given in Eq. (50).

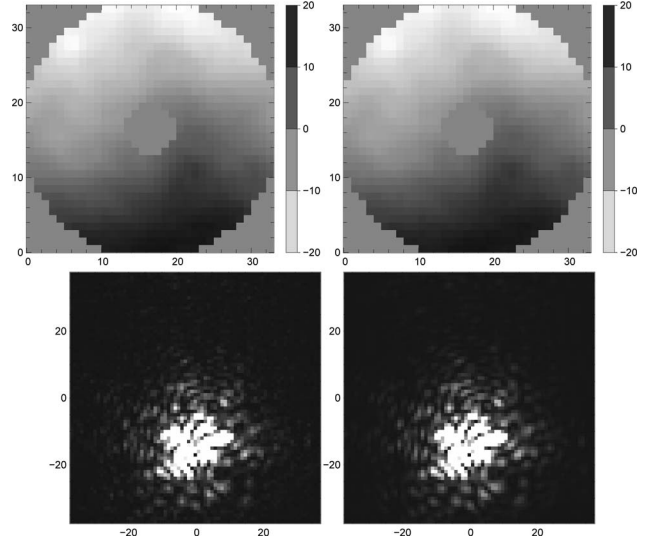


Fig. 1. Example of phase screens (left) and speckle images (right). The top panels show the true wavefront and the corresponding simulated image; the bottom panels show the recovered phase and corresponding model image. The phase scales are given in radians. The conditions of the simulation were $D/r_0 = 11$ and $N_{\text{ph}} = 50,000$ photons.

= 50,000 photons per image and different D/r_0 from 5 to 11 in order to test the algorithm performance regarding the turbulence strength. The SNR per speckle is from $\text{SNR} \approx 45$ to 22, respectively, for such conditions. Then we have run simulations with the same $D/r_0 = 8$ and different mean counts per image from $N_{\text{ph}} = 50,000$ to 1500 in order to see the influence of the source brightness alone. This corresponds to $\text{SNR} \approx 28$ to 5, respectively. Finally, we have performed a last test with the same $D/r_0 = 11$ and the same mean count per image $N_{\text{ph}} = 50,000$, but with two different samplings of the telescope field of view: 33×33 or 65×65 pixels for the wavefronts and with 75×75 or 135×135 pixels for the images. Indeed, this is of particular importance regarding the modulo- 2π phase ambiguity, because the smoothness brought by the prior and by the reparametrization depends on the sampling (the more the sampling, the more the smoothness and the less the ambiguity). All the different considered conditions of our simulations are detailed in Table 1. Typical simulated and retrieved phases are plotted in Fig. 1 with their corresponding images.

In order to estimate the typical number of random starts required by our global optimization strategy, we

proceed as follows: For each simulated image among the 250 images considered for each noise and turbulence level, we purposely attempt 200 independent starts, each followed by 70 local optimization steps. We record the best obtained error with respect to the number of trials for each image, regardless of whether the likelihood threshold has been reached or not. We then compute the average and the first, second (median), and third quartiles of the corresponding distribution of the best errors with respect to the number of random starts. Figure 2 shows these statistics for the different conditions we considered. These curves show that, although no restarts are required for the easiest cases (e.g., $D/r_0 \approx 5$), the number of trials re-

quired to achieve a given accuracy, say, 1 rad rms, clearly augments as the conditions worsen. This is expected partly because for higher D/r_0 ratios, the phase variance and hence the modulo- 2π ambiguities are more important, which results in much more local minima in the penalty criterion to be optimized. As we can see in Fig. 2, the sampling is critical regarding these ambiguities. Indeed, we can see that the number of trials to reach 1 rad rms is three times smaller when the model is oversampled. This is because the phase samples are less subject to the modulo- 2π wrapping ambiguity with proper sampling. As we can also see, the number of required restarts is less sensitive to the noise level than to the turbulence

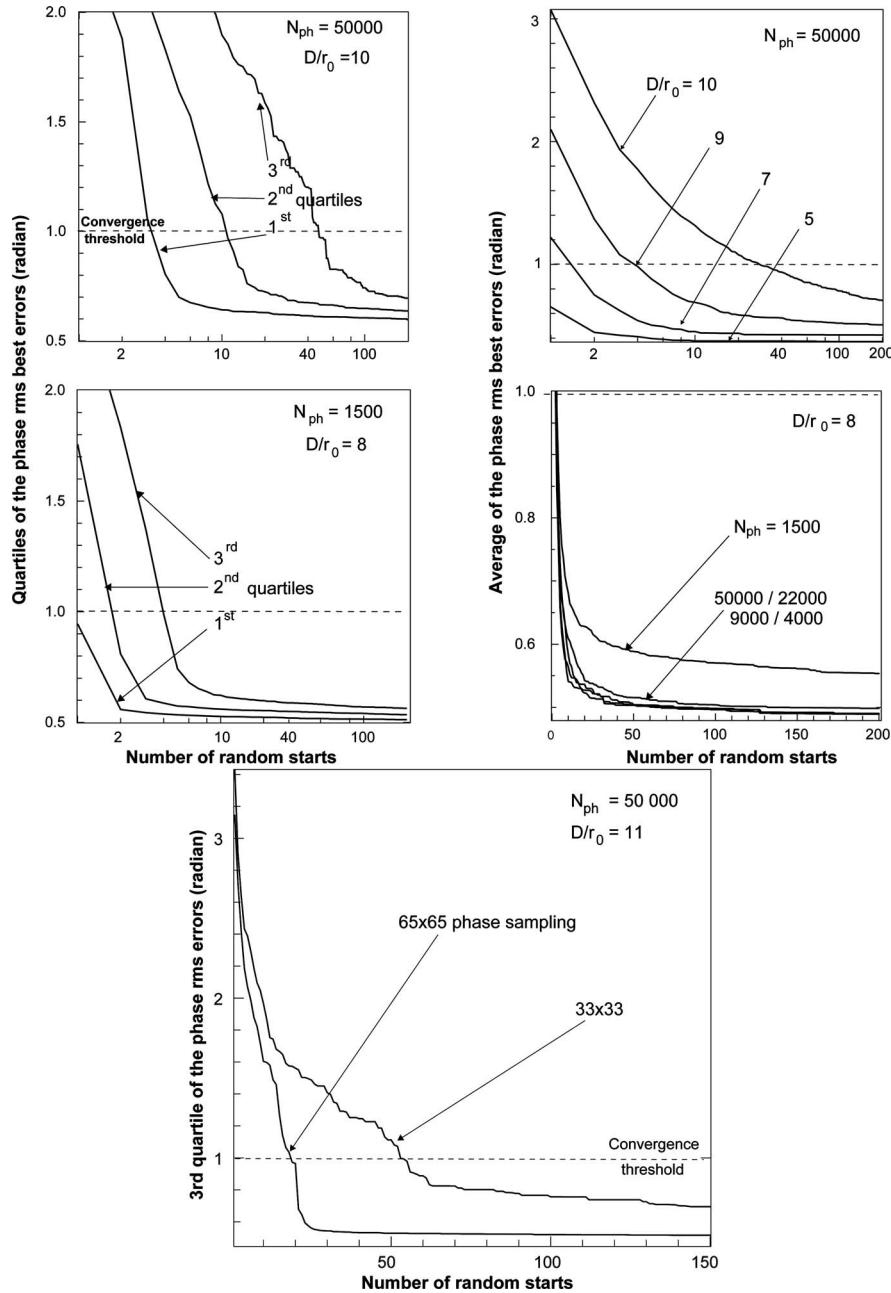


Fig. 2. Quartiles and average for uncorrelated wavefronts of the best phase rms errors achieved with respect to the number of random phase starts. Different turbulence conditions and mean photon counts are considered. The bottom panel shows the importance of model sampling for the success of the optimization for the same noise and turbulence levels. Indeed, finer sampling of the model enables better prior and reparametrization smoothness regarding the modulo- 2π phase wrapping ambiguity.

strength. In all the cases a modest number of random starts is sufficient with our algorithm: We find that fewer than 17 restarts are needed to solve 75% of the phase retrieval problems for a turbulence strength as high as $D/r_0=11$ and a mean flux as low as $N_{ph}=1500$ photons per image.

The cumulative distributions of the final rms errors between the retrieved and the true phases, in Fig. 3, assess the final accuracy of our global optimization. To compute these distributions, we have applied the local algorithm until full convergence, for the phase associated with the best adequation criterion f_{ML} , after the 200 random trials for each image. The curves show that the proposed method is very effective, even for the most difficult cases of strong turbulence and/or low photon counts. The limitation of 200 random starts explains that about 15% of the reconstructions fail at 1500 photons per image with $D/r_0=8$. Indeed, we have checked that this failure can be completely avoided if a higher limit is set on the maximum number of restarts.

Finally, the correlation between the values of the likelihood criterion f_{ML} and the phase errors, after the random part of the global optimization and before the full local optimization, is represented in Fig. 4. The distribution of the phase errors with respect to the phase quality, as

measured by the likelihood term f_{ML} , is indeed critical as f_{ML} is the only value that can be computed in practice. The figures show that all the recovered phases with $f_{ML} \leq f_{ML}^{min}$ have a rms error less than 1 rad over the pupil. This validates our criterion of initial phase selection on the basis of its likelihood after a few local optimization steps.

C. Time-Correlated Wavefronts

For correlated successive wavefronts, our strategy is to run only the local optimization part of our algorithm and to use the previous retrieved phase as the initial solution of the local optimization. Following Taylor's assumption for the slow temporal evolution of the wavefronts, we have simulated time-correlated wavefronts by translating a large Kolmogorov aberrated wavefront over our pupil. As advocated by Roddier *et al.* [22], the translation of a single aberrated wavefront cannot really mimic the boiling observed for true wavefronts. We have also checked our method with two translating layers to simulate the turbulence boiling. We have observed that the results are essentially the same and that the usefulness of the strategy does not depend on the details of the phase changes between successive realizations. To simplify the analysis

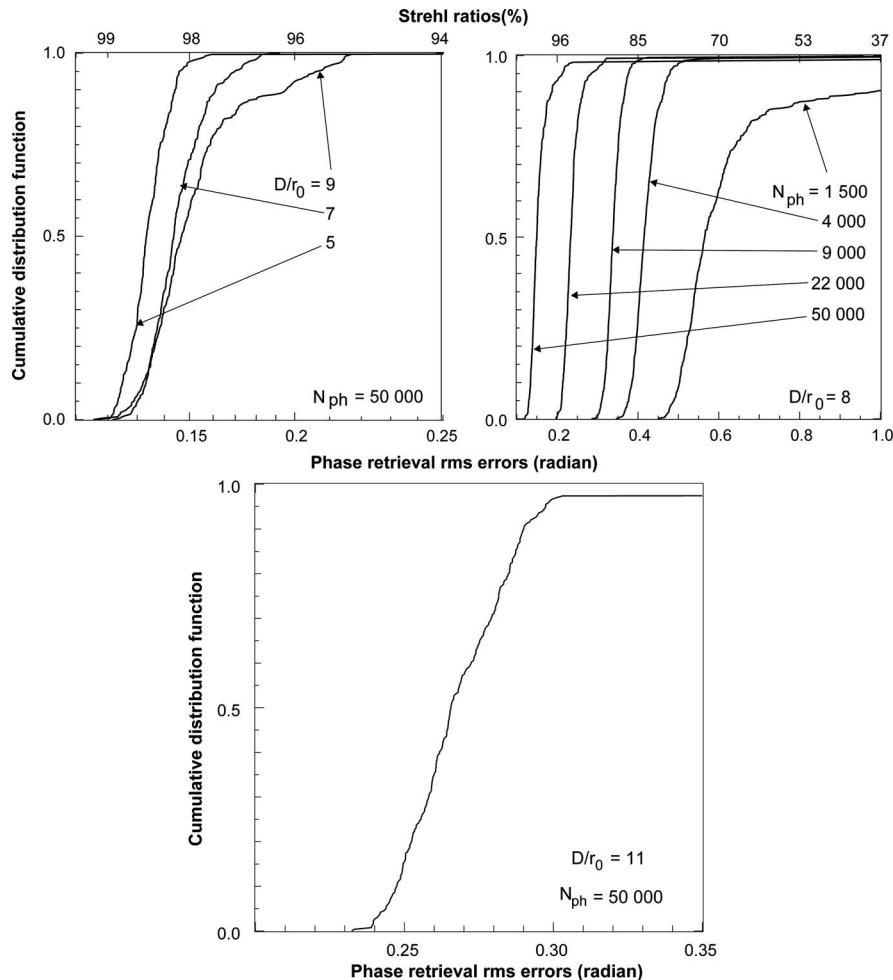


Fig. 3. Cumulative distribution functions $\Pr f(x \leq X)$ of the phase retrieval rms errors, in radian units, for different turbulence and noise conditions.

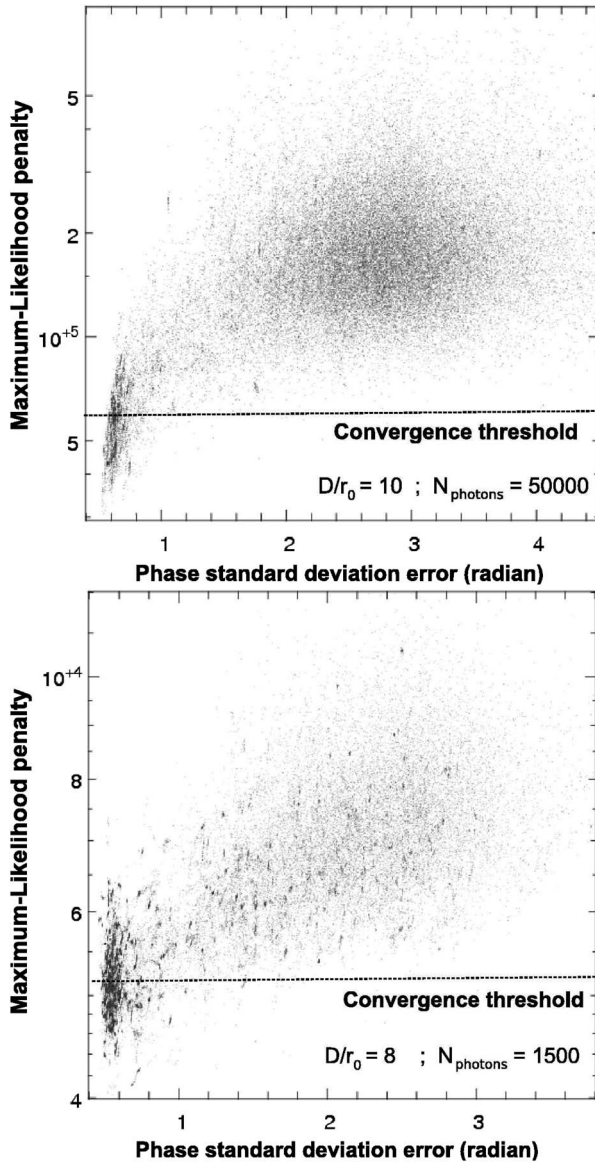


Fig. 4. Values of the ML penalty after 70 iterations of the local optimization algorithm versus the phase standard deviation error (rms error) with respect to the true phase. Different turbulence conditions and mean photon counts are considered. The threshold likelihood $f_{\text{ML}}^{\text{min}}$ is indicated for each case. The correlation between the best (smallest) penalties and the smallest phase errors is clearly seen.

of our strategy for correlated successive wavefronts, the results presented here were therefore obtained with a single translating layer.

We have considered different wind speeds, resulting in translations of $\Delta x = r_0/10$, $\Delta x = r_0/3$, and $\Delta x = r_0$ between two successive wavefronts. The other parameters of the simulations are a turbulence of $D/r_0 = 11$ and a mean flux of $N_{\text{ph}} = 10^5$ photons per image. The cumulative distribution function of the errors for $\Delta x = r_0$ is reported in Fig. 5. No phase retrieval fails, so the proposed strategy is effective for correlated wavefronts. We also note that the ambiguity between $\varphi(\mathbf{r})$ and $-\varphi(-\mathbf{r})$ no longer exists, providing the first retrieved phase of the sequence is correct with respect to that point.

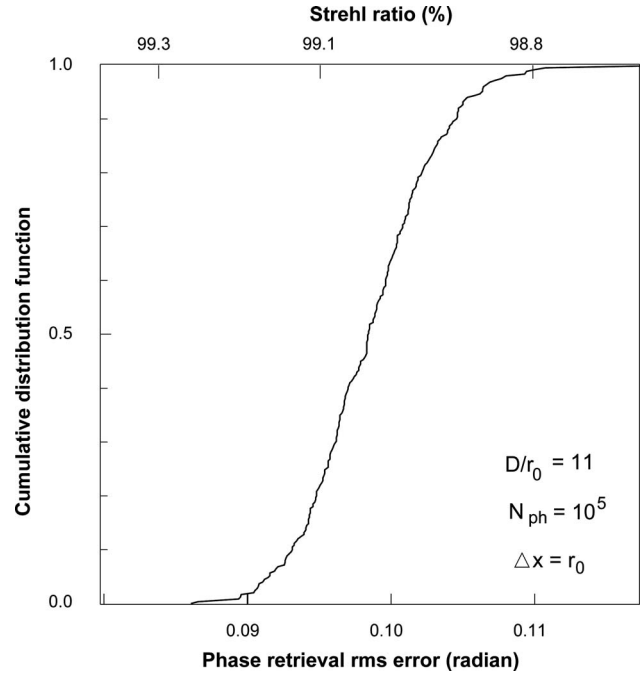


Fig. 5. Cumulative distribution functions of the phase retrieval rms errors (radian units) for time-correlated wavefronts.

7. DISCUSSION

We have presented a method to solve the nonlinear problem of phase retrieval from a single short-exposure monochromatic focused image. We have shown that the MAP approach, with phase smoothness prior, plays a deciding role in disentangling most of the ambiguities of the problem. Indeed, we observe that the global MAP solution always yields an accurate phase estimate.

We have based the global optimization required to obtain this solution on an efficient local NNLS method. In our local optimization algorithm, phase ambiguities and local minima are attenuated by the combination of (i) adaptive tuning of the smoothness level, (ii) appropriate phase parametrization, and (iii) tip-tilt bounding. Additional difficulties arise because of the nonlinearities of the MAP criterion and because of the large number of parameters. However, we achieve fast local convergence thanks to a subspace formulation of the optimization problem and a careful reordering of the computation of the partial derivatives.

For uncorrelated wavefronts, we have proposed a fast global optimization strategy based on random starts with early termination triggered by the likelihood of the sought phase. This strategy has proven practical efficiency: In more than 75% of the cases and up to $D/r_0 \sim 11$, fewer than 17 different starts are necessary to find the global optimum. It is worth noting that, owing to its intrinsic flexibility, our algorithm can cope with completely different combinations of difficulties, depending on the particular turbulence strength and noise level. The worst turbulence condition we have dealt with corresponds to ~ 120 deg of freedom for an adaptive optics system. This is a major advance compared with the ~ 16 deg of freedom previously achieved with a single image [4] and owing to the increase of complexity with D/r_0 .

In the case of time-correlated wavefronts, which would arise with continuous observations, we have found that the random start strategy can be completely avoided by using the previous retrieved phase. We are currently investigating means to avoid the random start strategy in more general cases (uncorrelated wavefronts).

Our approach can be improved to consider worse turbulence conditions. For example improvement of the tip-tilt bounding is expected to enable us to tackle higher D/r_0 . For even higher D/r_0 , we use polychromatic diversity to solve the arising phase wrapping ambiguities. These developments are the subjects of a forthcoming paper.

APPENDIX A: NONLINEAR LEAST SQUARES

This appendix details the expressions of the gradient and the Hessian that are needed to estimate the Newton step in the NNLS case. The penalty function $f = f_{\text{ML}} + \mu f_{\text{prior}}$ considered in this paper is a sum of terms with the general form

$$f(\boldsymbol{\psi}) = \frac{1}{2} \mathbf{e}(\boldsymbol{\psi})^T \cdot \mathbf{W} \cdot \mathbf{e}(\boldsymbol{\psi}), \quad (\text{A1})$$

where \mathbf{W} is a weighting (symmetrical positive semidefinite) matrix and the so-called vector of *residuals* $\mathbf{e}(\boldsymbol{\psi})$ is a function of the parameters $\boldsymbol{\psi}$. Indeed, in our case we have

$$\mathbf{e}_{\text{ML}}(\boldsymbol{\psi}) = \mathbf{d} - \mathbf{m}(\boldsymbol{\psi}); \quad \mathbf{W}_{\text{ML}} = \langle \mathbf{n} \cdot \mathbf{n}^T \rangle^{-1},$$

$$\mathbf{e}_{\text{prior}}(\boldsymbol{\psi}) = \boldsymbol{\psi}; \quad \mathbf{W}_{\text{prior}} = \mathbf{K}^T \cdot \mathbf{C}_{\varphi}^{-1} \cdot \mathbf{K}. \quad (\text{A2})$$

Considering the NLLS penalty in Eq. (A1), the gradient and the Hessian can be estimated by [14]

$$\mathbf{g}(\boldsymbol{\psi}) \stackrel{\text{def}}{=} \nabla f(\boldsymbol{\psi}) = \mathbf{J}(\boldsymbol{\psi})^T \cdot \mathbf{W} \cdot \mathbf{e}(\boldsymbol{\psi}), \quad (\text{A3})$$

$$\mathbf{H}(\boldsymbol{\psi}) \stackrel{\text{def}}{=} \nabla^2 f(\boldsymbol{\psi}) \simeq \mathbf{J}(\boldsymbol{\psi})^T \cdot \mathbf{W} \cdot \mathbf{J}(\boldsymbol{\psi}), \quad (\text{A4})$$

where $\mathbf{J}(\boldsymbol{\psi}) \stackrel{\text{def}}{=} \nabla \mathbf{e}(\boldsymbol{\psi})$ is the matrix of partial derivatives of the residuals ($J_{j,k}(\boldsymbol{\psi}) \stackrel{\text{def}}{=} [\partial e_j(\boldsymbol{\psi}) / \partial \psi_k]$). Note that the Hessian approximation in Eq. (A4) avoids the computation of the second-order partial derivatives of the residuals. This approximation is perfectly legitimate providing either (i) the residuals are not too nonlinear with respect to the parameters, (ii) the residuals are small, or (iii) the residuals are randomly distributed around zero [14,23]. Nonetheless, even if approximation in Eq. (A4) is crude, the trust region strategy (see Section 5) warrants that a suitable reduction step be taken.

APPENDIX B: FAST COMPUTATION BY REORDERING OF OPERATORS

This Appendix details how we compute the partial derivatives and the line search directions. To compute the ML partial derivatives, we need to apply the Jacobian \mathbf{J}_{ML} or its transposed matrix to a vector \mathbf{q} . Indeed, we have

$$\mathbf{g}_{\text{ML}} = \mathbf{J}_{\text{ML}}^T \cdot (\mathbf{W}_{\text{ML}} \cdot (\mathbf{d} - \mathbf{m})), \quad (\text{B1})$$

$$\mathbf{A}_{\ell,j}^{\text{ML}} = (\mathbf{J}_{\text{ML}} \cdot \mathbf{s}_{\ell})^T \cdot \mathbf{W}_{\text{ML}} \cdot (\mathbf{J}_{\text{ML}} \cdot \mathbf{s}_j). \quad (\text{B2})$$

It is also required when computing the line search directions:

$$\mathbf{s}_{3,4} = -\mu \times \mathbf{H}_{\text{prior}}^{-1} \cdot \mathbf{J}_{\text{ML}}^T \cdot (\mathbf{J}_{\text{ML}} \cdot \mathbf{s}_{1,3}). \quad (\text{B3})$$

Using the expression of the Jacobian,

$$\begin{aligned} \mathbf{J}_{j,\ell}^{\text{ML}} &= -\frac{\partial m_j(\boldsymbol{\psi})}{\partial p_{\ell}}, \\ &= 2\alpha_{\lambda} \text{Im} \left(a_{\lambda,j}^* \sum_k F_{j,k} A_{\lambda,k} \frac{\partial \varphi_{\lambda,k}}{\partial p_{\ell}} \right), \end{aligned} \quad (\text{B4})$$

and after reordering of the sum operator for efficient computation by means of a fast Fourier transform (FFT), we find

$$(\mathbf{J}_{\text{ML}} \cdot \mathbf{q})_j = 2\alpha_{\lambda} \text{Im} \left(a_{\lambda,j}^* \sum_k F_{j,k} A_{\lambda,k} \sum_{\ell} \frac{\partial \varphi_{\lambda,k}}{\partial p_{\ell}} q_{\ell} \right), \quad (\text{B5})$$

$$(\mathbf{J}_{\text{ML}}^T \cdot \mathbf{q})_{\ell} = 2\alpha_{\lambda} \sum_k \text{Im} \left(A_{\lambda,k} \sum_j F_{k,j} a_{\lambda,j}^* q_j \right) \frac{\partial \varphi_{\lambda,k}}{\partial p_{\ell}}, \quad (\text{B6})$$

where $\partial \varphi_{\lambda,k} / \partial p_{\ell} = K_{\ell,k}$.

ACKNOWLEDGMENTS

The development of our algorithm and all the numerical simulations and data processing have been done with Yorick, an interpreted programming language, which is freely available at <http://yorick.sourceforge.net/index.php>. We also want to thank the reviewers for their constructive comments and suggestions.

REFERENCES

1. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik (Stuttgart)* **35**, 237–246 (1972).
2. R. A. Gonsalves, "Phase retrieval from modulus data," *J. Opt. Soc. Am.* **66**, 961–964 (1976).
3. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
4. R. Irwin and R. G. Lane, "Phase retrieval with prior information," *J. Opt. Soc. Am. A* **15**, 2302–2311 (1998).
5. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Lett.* **21**, 829–832 (1982).
6. A. Tokovinin and S. Heathcote, "Donut: measuring optical aberrations from a single extrafocal image," *Publ. Astron. Soc. Pac.* **118**, 1165–1175 (2006).
7. F. Roddier, "The effects of atmospheric turbulence in optical astronomy," in *Progress in Optics*, E. Wolf, ed. (North Holland, 1981), pp. 281–376.
8. R. G. Lane and M. Tallon, "Wave-front reconstruction using a Shack–Hartmann sensor," *Appl. Opt.* **31**, 6902–6906 (1992).
9. E. Thiébaud and J.-M. Conan, "Strict *a priori* constraints for maximum-likelihood blind deconvolution," *J. Opt. Soc. Am. A* **12**, 485–492 (1995).

10. E. Thiébaud, "Introduction to image reconstruction and inverse problems," in *Optics in Astrophysics*, R. Foy and F.-C. Foy, eds. (Kluwer Academic, 2005), pp. 397–421.
11. R. G. Lane, "Methods for maximum-likelihood deconvolution," *J. Opt. Soc. Am. A* **13**, 1992–1998 (1996).
12. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, 1991).
13. E. P. Wallner, "Optimal wave-front correction using slope measurements," *J. Opt. Soc. Am.* **73**, 1771–1776 (1983).
14. J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer-Verlag, 2006).
15. J. J. Moré and D. C. Sorensen, "Computing a trust region step," *SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput.* **4**, 553–572 (1983).
16. J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Math. Comput.* **35**, 773–782 (1980).
17. E. Thiébaud, "Optimization issues in blind deconvolution algorithms," *Proc. SPIE* **4847**, 174–183 (2002).
18. J. Skilling and R. K. Bryan, "Maximum entropy image reconstruction: general algorithm," *Mon. Not. R. Astron. Soc.* **211**, 111–124 (1984).
19. C. Pichon and E. Thiébaud, "Non-parametric reconstruction of distribution functions from observed galactic discs," *Mon. Not. R. Astron. Soc.* **301**, 419–434 (1998).
20. C. Béchet, M. Tallon, and E. Thiébaud, "FRIM: minimum-variance reconstructor with a fractal iterative method," *Proc. SPIE* **6272**, 62722U (2006).
21. R. G. Lane, A. Glindemann, and J. C. Dainty, "Simulation of a Kolmogorov phase screen," *Waves Random Media* **2**, 209–224 (1992).
22. F. Roddier, J. M. Gilli, and G. Lund, "On the origin of speckle boiling and its effects in stellar speckle interferometry," *J. Opt.* **13**, 263–271 (1982).
23. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge U. Press, 1992).

Polychromatic phase retrieval with Kolmogorov self-adapting prior constraints

X. Rondeau^a, E. Thiébaud^a and M. Tallon^a

^aCRAL UMR5574 CNRS / ENS Lyon / UCBL, Observatoire de Lyon, 9 avenue Charles André,
69531 Saint Genis Laval Cedex, France

ABSTRACT

Phase retrieval is a very promising approach for wavefront sensing in the focal plane of ground-based large telescopes. It is a non-linear problem that must be solved by means of global optimization. Currently only multi-focal phase diversity algorithms are used in adaptive optics. They enable the correction of static aberrations. For speckle imaging the problem is increasingly multi-modal with the ratio D/r_0 . Yet thanks to an iterative Newton algorithm with self-adapting Kolmogorov prior information, we show from consistent modeling and simulations, that we could efficiently sense short exposure wavefronts at high D/r_0 from a single focal plane. We show that using data at different wavelengths with a proper polychromatic model would even enforce the convergence, thus making it an envisageable method to sense the returned flux of a polychromatic laser guide star (PLGS). For instance, we show that if we suppose the PLGS is not resolved, phase retrieval would enable an improvement in the centroid estimation in agreement with the Cramer-Rao lower bound. As a post-processing technique, our algorithm already has numerous potential applications for astronomy and for other domains. Thanks to the improvement of computing workstations and the optimization of the algorithm, applications involving real-time wavefront corrections should be soon possible.

Keywords: phase retrieval, wavefront sensing, speckle images, global optimization, prior constraints, tip-tilt, centroiding, polychromatic laser guide star

1. INTRODUCTION

In the framework of the Polychromatic Laser Guide Star for Adaptive Optics (ELP-OA), we have developed an algorithm to sense aberrated wavefronts from short exposure images at different wavelengths. ELP-OA aims at the correction of the atmospheric tip-tilt from chromatic differential measurements. That's why we use spectral diversity instead of the usual focal diversity to enforce our phase retrieval. Our algorithm framework is wider than the PLGS, in that we sense aberrated wavefronts very accurately even with a single wavelength. The paper is divided into four sections. First we recall the general phase retrieval model and extend it to a polychromatic model. Then we derive the properties of our global optimization algorithm. In section three we present the performances from consistent simulations, using one or several wavelengths. Section four is dedicated to the application of our algorithm to the tilt estimation for the PLGS. Finally we conclude on the future developments to make our phase retrieval approach real-world applicable.

2. MODELIZATION

2.1. Physics background

We first recall the continuous physics equations concerning short exposure images. We denote $\varphi_\lambda(\mathbf{r})$ the monochromatic aberrated phase on the pupil of a ground-based telescope at wavelength λ . Considering the atmosphere as an achromatic medium stratified in refractive index n , the phase at wavelength λ and at a position \mathbf{r} of the pupil plane writes:

$$\varphi_\lambda(\mathbf{r}) = \frac{2\pi}{\lambda} \delta(\mathbf{r})$$

where $\delta(\mathbf{r}) = \text{cste} + \int_0^h n(\mathbf{r}, z) dz$ is the optical path length through the transfer medium. Following the prescription used in the polychromatic laser guide star theory,¹ we only consider the wavelength dependency of the refractive index for the

E-mail: xrondeau@obs.univ-lyon1.fr, Phone: +33 (0)4-78-86-85-47

first orders (tip-tilt) of the Zernike modal decomposition of the aberrated phase. Denoting $\mathbf{b}(\mathbf{r})$ the tip-tilt functions in the Zernike basis and θ_λ the corresponding coefficients (in units of image plane shift) at wavelength λ , the chromatic phase writes:

$$\varphi_\lambda(\mathbf{r}) = \frac{\lambda_0}{\lambda} \phi(\mathbf{r}) + \theta_\lambda^\top \cdot \mathbf{b}(\mathbf{r}/\lambda) \quad (1)$$

where $\phi(\mathbf{r})$ is the phase at a reference wavelength λ_0 with piston and tip-tilt removed. If the pupil is symmetrical and centered at $\mathbf{r} = (0, 0)$, the tip-tilt functions simplify to:

$$\mathbf{b}(\mathbf{r}/\lambda) = \frac{2\pi}{\lambda} \mathbf{r}.$$

Neglecting the scintillation effects of the turbulence, the complex amplitude in the pupil at wavelength λ for a point-like source writes:

$$A_\lambda(\mathbf{r}) = P(\mathbf{r}) \exp(i\varphi_\lambda(\mathbf{r})) \quad (2)$$

where $P(\mathbf{r})$ is the (complex) amplitude transmission of the telescope, the so-called *pupil function*. Given the complex amplitude in the pupil, the instantaneous monochromatic brightness distribution in the direction \mathbf{x} writes:²

$$I_\lambda(\mathbf{x}) = \alpha_\lambda \left| \hat{A}_\lambda(\mathbf{x}) \right|^2 \quad (3)$$

where $\alpha_\lambda > 0$ accounts for the monochromatic flux of the source and a necessary normalization factor and where $\hat{A}_\lambda(\mathbf{x})$ is the Fourier transform of the complex amplitude in the pupil:

$$\hat{A}_\lambda(\mathbf{x}) = \iint A_\lambda(\lambda \mathbf{u}) e^{+2i\pi \mathbf{x} \cdot \mathbf{u}} d^2\mathbf{u} = \frac{1}{\lambda^2} \iint A_\lambda(\mathbf{r}) e^{+\frac{2i\pi}{\lambda} \mathbf{x} \cdot \mathbf{r}} d^2\mathbf{r} \quad (4)$$

where $\mathbf{u} = \mathbf{r}/\lambda$ is the spatial frequency. As we can see from Eqs. (3-4), the relation between the speckle image and the aberrated phase is strongly non-linear; this gives rise to most of the difficulties of the phase retrieval problem.

2.2. Discretization

Image detection by a CCD involves sampling of the brightness distribution. Neglecting the filtering by the pixels, the discrete model \mathbf{m}_λ of the monochromatic image writes:

$$m_{\lambda,j} = I_\lambda(\mathbf{x}_j) \quad (5)$$

where \mathbf{x}_j is the viewing direction of the j -th CCD pixel. Taking $\hat{A}_{\lambda,j} \equiv \hat{A}_\lambda(\mathbf{x}_j)$, the discretization is continued as follows:

$$m_{\lambda,j} = \alpha_\lambda \left| \hat{A}_{\lambda,j} \right|^2 = \alpha_\lambda \left| \sum_k F_{j,k} A_{\lambda,k} \right|^2 \quad (6)$$

where $F_{j,k} = \exp(2i\pi j k / N_{\text{pix}})$ is the discrete Fourier transform matrix (N_{pix} is the number of pixels along a dimension) and the discretized complex amplitude in the pupil plane is:

$$A_{\lambda,k} \equiv A_\lambda(\lambda \mathbf{u}_k) = P_{\lambda,k} \exp[i\varphi_{\lambda,k}] \quad (7)$$

where \mathbf{u}_k is the k -th discrete spatial frequency, $P_{\lambda,k} \equiv P(\lambda \mathbf{u}_k)$ and $\varphi_{\lambda,k} \equiv \varphi_\lambda(\lambda \mathbf{u}_k)$ are the sampled pupil function and chromatic phase. The sampled phase further expands to:

$$\begin{aligned} \varphi_{\lambda,k} \equiv \varphi_\lambda(\lambda \mathbf{u}_k) &= \frac{\lambda_0}{\lambda} \phi(\lambda \mathbf{u}_k) + \theta_\lambda^\top \cdot \mathbf{b}(\mathbf{u}_k) \\ &= \sum_\ell L_{\lambda,k,\ell} \phi_\ell + \theta_\lambda^\top \cdot \mathbf{b}_k \end{aligned} \quad (8)$$

where $\mathbf{b}_k \equiv \mathbf{b}(\mathbf{u}_k)$, and where \mathbf{L}_λ is a linear interpolating factor (which accounts for the phase and frequency scaling by λ_0/λ), and where ϕ is the discretized reference phase (without piston and tip-tilt):

$$\phi_\ell \equiv \phi(\lambda_0 \mathbf{u}_\ell). \quad (9)$$

In the following we may introduce an ad-hoc basis of functions (so-called *phase modes*) \mathbf{K} to represent the phase, then:

$$\phi = \mathbf{K} \cdot \psi \quad (10)$$

in other words: $\phi_\ell = \sum_m K_{\ell,m} \psi_m$ where ψ is the vector of mode weights.

2.3. Noise and model errors

A consistent model for real data must at least account for the photon and the CCD noises. After a CCD dark subtraction meant to take off the CCD electronic bias, the data $\mathbf{d} = \{\mathbf{d}_{\lambda_1}, \mathbf{d}_{\lambda_2}, \dots\}$ is related to the discrete model $\mathbf{m} = \{\mathbf{m}_{\lambda_1}, \mathbf{m}_{\lambda_2}, \dots\}$ by:

$$\mathbf{d}_\lambda = \mathbf{m}_\lambda + \mathbf{n}_\lambda \quad (11)$$

where $\mathbf{n} = \{\mathbf{n}_{\lambda_1}, \mathbf{n}_{\lambda_2}, \dots\}$ accounts for modelization errors and noise contributions: photon-noise of the signal (which follows Poissonian statistics) plus CCD dark current and read-out noise (with approximately Gaussian statistics).

3. PHASE RESTORATION ALGORITHM

The principle of the algorithm comes from Pichon and Thiébaud³ and is transposed to the case of phase retrieval. We first introduce the phase retrieval settings such as the regularization, the parametrization and the derivatives before recalling the principle of the algorithm.

3.1. Bayesian approach

Taking into consideration the statistic of the errors \mathbf{n} from a maximum likelihood (ML) point of view, leads us to seek for the phase parameters \mathbf{p}_{ML} that maximize the probability of the data given the model:

$$\begin{aligned} \mathbf{p}_{\text{ML}} &= \arg \max_{\mathbf{p}} \Pr(\mathbf{d} \mid \mathbf{m}(\mathbf{p})) \\ &= \arg \min_{\mathbf{p}} f_{\text{ML}}(\mathbf{p}) \end{aligned} \quad (12)$$

where the model \mathbf{m} is a function of the parameters $\mathbf{p} = \{\phi, \theta_{\lambda_1}, \theta_{\lambda_2}, \dots\}$ or $\mathbf{p} = \{\psi, \theta_{\lambda_1}, \theta_{\lambda_2}, \dots\}$ and $f_{\text{ML}}(\mathbf{p}) = -\log \Pr(\mathbf{d} \mid \mathbf{m}(\mathbf{p}))$. This approach however neither smoothes out the noise,⁴ nor is robust with respect to the ill-conditioning of the inverse problem, nor helps to avoid degeneracies.

Staying in the Bayesian framework, a better approach is to seek for the model parameters which maximize the likelihood of the model given the data. This naturally leads us to consider *a priori constraints* for the parameters by means of their expected statistics (e.g., Kolmogorov law for the phase model). The resulting maximum a posteriori (MAP) solution writes:

$$\begin{aligned} \mathbf{p}_{\text{MAP}} &= \arg \max_{\mathbf{p}} \Pr(\mathbf{d} \mid \mathbf{m}(\mathbf{p})) \Pr(\mathbf{m}(\mathbf{p})) \\ &= \arg \min_{\mathbf{p}} [f_{\text{ML}}(\mathbf{p}) + \mu f_{\text{prior}}(\mathbf{p})] \end{aligned} \quad (13)$$

where $\mu f_{\text{prior}}(\mathbf{p}) = -\log [\Pr(\mathbf{m}(\mathbf{p}))]$ is the neg-log likelihood of the prior. In inverse problem jargon, f_{prior} is the regularization term and μ is an hyperparameter required to tune the relative strength of the a priori constraints with respect to the data constraints. Alternatively, μ can be seen as a Lagrange multiplier tuned so as to match a constraint about the expected value of f_{ML} or f_{prior} . For instance, following Gull,⁵ one would choose μ so that $f_{\text{ML}}(\mathbf{p}_{\text{MAP}}) + \mu f_{\text{prior}}(\mathbf{p}_{\text{MAP}}) = \frac{1}{2} N_{\text{data}}$, where N_{data} is the number of measurements. As detailed below, we have implemented this kind of requirement by appropriately tuning μ at each iteration of our minimization algorithm.

3.2. Maximum-likelihood term

When having both Poisson and Gaussian noise affecting the data, taking Gaussian statistics is a robust approach.⁶ Under this approximation, the ML penalty term writes:

$$f_{\text{ML}} = \frac{1}{2} (\mathbf{d} - \mathbf{m})^\top \cdot \mathbf{W} \cdot (\mathbf{d} - \mathbf{m}) \quad (14)$$

where $\mathbf{W} = \mathbf{C}_n^{-1}$ is the inverse of the covariance matrix $\mathbf{C}_n = \langle \mathbf{n} \cdot \mathbf{n}^\top \rangle$ of the errors \mathbf{n} .

Since the model is non-linear, f_{ML} is not a quadratic function of the phase parameters and there is no direct solution for phase retrieval. The minimization of f_{ML} must be iterated with successive inversions along local quadratic approximations of f_{ML} . As pointed out by Irwan & Lane,⁷ the image model is unchanged neither when a piston is added to the phase,

nor when $2k\pi$ is added to any phase value, nor when $\varphi(\mathbf{r})$ is replaced by $-\varphi(-\mathbf{r})$ providing the pupil is symmetric. The two first points constitute the phase unwrapping problem. To overcome unwrapping Michelson and Benoit⁸ used several wavelengths when measuring the meter. Fitting several wavelengths in the algorithm may therefore also restrict the acceptable solution with respect to unwrapping. Accounting for regularization (as explained below) is also expected to solve for the unwrapping since regularization enforces the smoothness (hence continuity) of the restored phase. The last point is more general, as there is a quasi exact degeneracy on the sign of each even mode of a modal phase decomposition. Again we can expect that such degeneracies can be disentangled by using a smoothness prior.

3.3. Regularization

Regularization is mandatory to cope with the ill-conditioned nature of most inverse problems.⁹ As previously advocated, regularization can also help to avoid local minima of the maximum-likelihood penalty alone and disentangle degeneracies of the phase restoration problem.

We have experienced that a simple Tikhonov regularization $\phi^\top \cdot \phi$ is not able to solve all the degeneracies of the phase restoration problem. Since the turbulent phase (after removal of the piston) is known to obey Gaussian statistics, the proper definition of the regularization term in the Bayesian framework is:

$$f_{\text{prior}} = \phi^\top \cdot \mathbf{C}_\phi^{-1} \cdot \phi \quad (15)$$

where $\mathbf{C}_\phi = \langle \phi \cdot \phi^\top \rangle$ is the covariance matrix of the phase ϕ .

As shown by Wallner,¹⁰ there is a simple expression that links the covariance matrix \mathbf{C}_φ of a piston-removed phase screen φ , to the Kolmogorov structure function

$$D_\varphi(\mathbf{r}_i, \mathbf{r}_j) \simeq 6.88 \times \left| \frac{\mathbf{r}_i - \mathbf{r}_j}{r_0} \right|^{\frac{5}{3}} \quad (16)$$

where r_0 is Fried's parameter. Phase retrieval being not meant at assessing the piston, we can follow this approach for the regularization term. Moreover, observing that the tilt is heavily weighting in the covariance, and as our model leads us into separating a chromatic tilt from an achromatic phase, we rather use the covariance of a Kolmogorov phase ϕ with the piston and the tilt being removed:

$$\begin{aligned} \mathbf{C}_\phi(\mathbf{r}_i, \mathbf{r}_j) = & -\frac{1}{2} D_\varphi(\mathbf{r}_i, \mathbf{r}_j) + \frac{1}{2} \sum_{\ell} D_\varphi(\mathbf{r}_\ell, \mathbf{r}_j) [P(\mathbf{r}_\ell) + c \mathbf{b}(\mathbf{r}_\ell)^\top \cdot \mathbf{b}(\mathbf{r}_i)] \\ & + \frac{1}{2} \sum_{\ell} D_\varphi(\mathbf{r}_i, \mathbf{r}_\ell) [P(\mathbf{r}_\ell) + c \mathbf{b}(\mathbf{r}_\ell)^\top \cdot \mathbf{b}(\mathbf{r}_j)] \\ & - \frac{1}{2} \sum_{k, \ell} D_\varphi(\mathbf{r}_k, \mathbf{r}_\ell) [P(\mathbf{r}_k) + c \mathbf{b}(\mathbf{r}_k)^\top \cdot \mathbf{b}(\mathbf{r}_i)] [P(\mathbf{r}_\ell) + c \mathbf{b}(\mathbf{r}_\ell)^\top \cdot \mathbf{b}(\mathbf{r}_j)] \end{aligned} \quad (17)$$

where the pupil P is normalized to have an unit integral and the coefficient c normalizes the tilt by the squared sum along one direction.

3.4. Phase parametrization

Since a change of variable results in a different condition number of the inverse problem, proper phase parametrization can be a deciding issue in the phase restoration problem. We have tested different function basis to represent the phase: Fourier decomposition, Gram-Schmidt orthonormalized basis of polynomials, and directly the sampled phase ϕ where $\phi_\ell = \phi(\lambda_0 \mathbf{u}_\ell)$. None of these parametrizations have brought any significant improvement. Our final choice was therefore motivated by the minimal number of parameters to account for the smallest structures in the image. As demonstrated by Lane and Tallon,¹¹ the Karhunen-Loève polynomials match this requirement.

In practice, we derive the Karhunen-Loève basis for our sampled phase ϕ by using the spectral decomposition or the singular value decomposition¹² (SVD) of its covariance matrix:

$$\mathbf{C}_\phi = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{U}^\top \quad (18)$$

where Σ is diagonal and \mathbf{U} is orthonormal. Taking $\psi = \Sigma^{-\frac{1}{2}} \cdot \mathbf{U}^\top \cdot \phi$ and conversely $\phi = \mathbf{K} \cdot \psi$ with $\mathbf{K} = \mathbf{U} \cdot \Sigma^{\frac{1}{2}}$ yields the Karhunen-Loève phase parameters ψ such that $\mathbf{C}_\psi = \mathbf{I}$ and for which the regularization in Eq. (15) simplifies to:

$$f_{\text{prior}} = \psi^\top \cdot \psi. \quad (19)$$

3.5. Non-linear least squares

Since all penalty functions f_{ML} (for all wavelengths) and f_{prior} are sum of squares, local optimization by our algorithm is designed after other non-linear least squares (NLLS) methods. The penalty function $f_{\text{ML}} + \mu f_{\text{prior}}$ considered in this paper is a sum of terms with general form:

$$f(\mathbf{p}) = \frac{1}{2} \mathbf{e}(\mathbf{p})^\top \cdot \mathbf{W} \cdot \mathbf{e}(\mathbf{p}) \quad (20)$$

where \mathbf{W} is a weighting (symmetric positive semi-definite) matrix and the so-called vector of *residuals* $\mathbf{e}(\mathbf{p})$ is a function of the parameters \mathbf{p} . Indeed, in the case of $f_{\text{ML}}(\mathbf{p})$, we have $\mathbf{e}_{\text{ML}}(\mathbf{p}) = \mathbf{d} - \mathbf{m}(\mathbf{p})$ and $\mathbf{W}_{\text{ML}} = \langle \mathbf{n} \cdot \mathbf{n}^\top \rangle^{-1}$. In the case of $f_{\text{prior}}(\mathbf{p})$, we have $\mathbf{e}_{\text{prior}}(\mathbf{p}) = \phi$ and $\mathbf{W}_{\text{prior}} = \langle \phi \cdot \phi^\top \rangle^{-1}$.

By Taylor expansion, the penalty change for a small variation $\delta\mathbf{p}$ of the parameters writes:

$$\begin{aligned} \delta f(\mathbf{p}, \delta\mathbf{p}) &\equiv f(\mathbf{p} + \delta\mathbf{p}) - f(\mathbf{p}) \\ &= \mathbf{g}(\mathbf{p})^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H}(\mathbf{p}) \cdot \delta\mathbf{p} + \mathcal{O}(\|\delta\mathbf{p}\|^3) \end{aligned} \quad (21)$$

where $\mathbf{g}(\mathbf{p}) \equiv \nabla f(\mathbf{p})$ and $\mathbf{H}(\mathbf{p}) \equiv \nabla \nabla f(\mathbf{p})$ are respectively the gradient (vector of partial first derivatives) and Hessian (matrix of partial second derivatives) of $f(\mathbf{p})$. Iterative minimization algorithms make use of Taylor-expansion in Eq. (21) to compute a *good* step $\delta\mathbf{p}$ and update the set of parameters by $\mathbf{p} \leftarrow \mathbf{p} + \delta\mathbf{p}$. For instance, the quadratic approximation:

$$\delta f_{\text{quad}}(\mathbf{p}, \delta\mathbf{p}) = \mathbf{g}(\mathbf{p})^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H}(\mathbf{p}) \cdot \delta\mathbf{p} \quad (22)$$

of the local behaviour of $f(\mathbf{p} + \delta\mathbf{p}) - f(\mathbf{p})$ can be used to derive the parameter step:

$$\delta\mathbf{p}_{\text{quad}}(\mathbf{p}) = -\mathbf{H}(\mathbf{p})^{-1} \cdot \mathbf{g}(\mathbf{p}) \quad (23)$$

providing $\delta\mathbf{p}_{\text{quad}}$ is small enough to neglect the non-quadratic terms in the Taylor expansion (21). A convenient way to insure that this constraint is satisfied is to add a step length penalty and consider the step:

$$\delta\mathbf{p}_\eta(\mathbf{p}) = -[\mathbf{H}(\mathbf{p}) + \eta \mathbf{I}]^{-1} \cdot \mathbf{g}(\mathbf{p}) \quad (24)$$

which minimizes $\delta f_{\text{quad}}(\mathbf{p}, \delta\mathbf{p}) + \frac{1}{2} \eta \|\delta\mathbf{p}\|^2$. In Eq. (24), \mathbf{I} is the identity matrix and $\eta \geq 0$ is a Lagrange multiplier introduced to account for the constraint that $\delta\mathbf{p}$ must be small. In our algorithm, we use a trust-region method¹³ to control the step length. Given the radius $\Delta > 0$ of the region for which the quadratic approximation is valid, the trust-region algorithm finds a value $\eta \geq 0$ such that $\|\delta\mathbf{p}_\eta(\mathbf{p})\| \leq \Delta$. The trust-region algorithm also takes care of updating properly the value of Δ .

Considering the NLLS penalty in Eq. (20), the gradient and Hessian (required to compute the parameter step) can be derived as follows:

$$\mathbf{g}(\mathbf{p}) \equiv \nabla f(\mathbf{p}) = \mathbf{J}(\mathbf{p})^\top \cdot \mathbf{W} \cdot \mathbf{e}(\mathbf{p}) \quad (25)$$

$$\mathbf{H}(\mathbf{p}) \equiv \nabla \nabla f(\mathbf{p}) \simeq \mathbf{J}(\mathbf{p})^\top \cdot \mathbf{W} \cdot \mathbf{J}(\mathbf{p}) \quad (26)$$

where $\mathbf{J}(\mathbf{p}) \equiv \nabla \mathbf{e}(\mathbf{p})$ is the matrix of partial derivatives of the residuals:

$$J_{j,k}(\mathbf{p}) = \frac{\partial e_j(\mathbf{p})}{\partial p_k}. \quad (27)$$

Note that the Hessian approximation in Eq. (26) avoids the computation of the second order partial derivatives of the residuals. This approximation can be done if (i) the residuals are not too non-linear with respect to the parameters, or (ii) the residuals are small, or (iii) the residuals are not too large and randomly distributed around zero.

3.6. Partial Derivatives

Our regularization term being exactly quadratic with respect to the parameters, the matrix of partial derivatives of the residuals is trivial to obtain: $\mathbf{J}_{\text{prior}} = \mathbf{I}$, *i.e.* the identity. If the phase is parameterized by ϕ , then $\mathbf{W}_{\text{prior}} = \mathbf{C}_{\phi}^{-1}$; otherwise the phase is parameterized by ψ , and $\mathbf{W}_{\text{prior}} = \mathbf{C}_{\psi}^{-1} = \mathbf{I}$.

For the maximum likelihood term, the expression of \mathbf{J}_{ML} is more complicated:

$$\begin{aligned} J_{\text{ML},\lambda,j,\ell} &= -\frac{\partial m_{\lambda,j}(\mathbf{p})}{\partial p_{\ell}} \\ &= 2\alpha_{\lambda} \text{Im} \left(\hat{A}_{\lambda,j}^* \sum_k F_{j,k} A_{\lambda,k} \frac{\partial \varphi_{\lambda,k}}{\partial p_{\ell}} \right). \end{aligned} \quad (28)$$

There are two kind of model parameters: tip-tilt and de-tilted phase. The partial derivatives of the chromatic phase with respect to the tip-tilt parameters write:

$$\frac{\partial \varphi_{\lambda,k}}{\partial \theta_{\lambda,\ell}} = b_{\ell,k} \quad (29)$$

where $\ell = 1$ or $\ell = 2$ since there are two possible tip-tilt directions. For instance, in the case of a symmetrical pupil, $\mathbf{b}_k = 2\pi \mathbf{u}_k$ where $\mathbf{u}_k = (u_k, v_k)$ is the k -th spatial frequency, hence: $b_{1,k} = 2\pi u_k$ and $b_{2,k} = 2\pi v_k$. The partial derivatives of the chromatic phase with respect to the parameters ϕ write:

$$\frac{\partial \varphi_{\lambda,k}}{\partial \phi_{\ell}} = L_{\lambda,k,\ell}. \quad (30)$$

When parameters ψ are used instead, we need:

$$\frac{\partial \varphi_{\lambda,k}}{\partial \psi_m} = \sum_{\ell} L_{\lambda,k,\ell} K_{\ell,m}. \quad (31)$$

In practice, once the proper expression for the partial derivatives of the phase model is plugged into Eq. (28), the action of \mathbf{J}_{ML} can be efficiently computed by means of a fast Fourier transform (FFT) after re-ordering of the sum operators.

3.7. Tuning of prior constraints

Our derivation of the criterion $f_{\text{ML}} + \mu f_{\text{prior}}$ is based on Bayesian considerations. In this framework, there should be no reason to tune the relative weight of the prior. This would however require to have good estimates of the statistics of the noise and of the turbulence. In practice, precise measurements of the Fried's parameter r_0 , of the CCD read-out noise, of the CCD gain, etc, may not be available. This justify the need to tune the value of μ in the algorithm.

In our case, another reason to tune the relative weight of the prior is to favor the convergence of the algorithm and to avoid small local minima. This is achieved by starting the algorithm with a high regularization level and by slowly reducing the value of μ as the iterations proceed. In that way, the regularization is kept to a level sufficient to disentangle some of the degeneracies of the maximum likelihood term f_{ML} alone. This kind of strategy has proven effectiveness for other inverse problems under maximum entropy regularization.^{3,14}

Our objective is to have after complete convergence of the algorithm: $f_{\text{ML}}(\mathbf{p}_{\mu}) + \mu f_{\text{prior}}(\mathbf{p}_{\mu}) = f_{\text{final}}$ where \mathbf{p}_{μ} is the set of parameters obtained with a regularization level μ and where f_{final} is the target value of the total penalty. For instance, following Gull's prescription,⁵ one would take $f_{\text{final}} = \frac{1}{2} N_{\text{data}}$. At every iteration of our algorithm, the parameter step depends on the values of η (which controls the step length *via* the trust-region radius Δ) and μ (which controls the regularization level) as:

$$\delta \mathbf{p}_{\mu} = -(\mathbf{H}_{\text{ML}} + \mu \mathbf{H}_{\text{prior}} + \eta \mathbf{I})^{-1} \cdot (\mathbf{g}_{\text{ML}} + \mu \mathbf{g}_{\text{prior}}) \quad \text{with } \eta \geq 0 \text{ such that: } \|\delta \mathbf{p}_{\mu}\| \leq \Delta \quad (32)$$

where \mathbf{g}_{ML} and \mathbf{H}_{ML} are the gradient and Hessian of $f_{\text{ML}}(\mathbf{p})$ and where $\mathbf{g}_{\text{prior}}$ and $\mathbf{H}_{\text{prior}}$ are the gradient and Hessian of $f_{\text{prior}}(\mathbf{p})$. The exact meaning of Eq. (32) is that, given the values of μ and Δ , the value of η is automatically computed by

the trust-region algorithm of Moré & Sorensen¹³ to ensure that $\|\delta\mathbf{p}_\mu\| \leq \Delta$. In other words, η is virtually a function of μ and Δ . For a given value of Δ , we make use of the quadratic approximations:

$$\delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}) \simeq \mathbf{g}_{\text{ML}}^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H}_{\text{ML}} \cdot \delta\mathbf{p} \quad (33)$$

$$\delta f_{\text{prior}}^{\text{quad}}(\delta\mathbf{p}) \simeq \mathbf{g}_{\text{prior}}^\top \cdot \delta\mathbf{p} + \frac{1}{2} \delta\mathbf{p}^\top \cdot \mathbf{H}_{\text{prior}} \cdot \delta\mathbf{p} \quad (34)$$

to anticipate the variation of each penalty terms if a step $\delta\mathbf{p}_\mu$ as given by Eq. (32) is taken. In order to vary the value of μ on the fly we modified our trust-region algorithm as follows:

1. While the model is in bad agreement with the data, *i.e.* while $f_{\text{ML}}(\mathbf{p}) \geq f_{\text{final}}$, the regularization level is adjusted so that the maximum likelihood term is reduced by a small fraction ε of what is best expected from local minimization. This warrants that the agreement with the data get improved while maintaining strong control by the prior constraints on the step effectively taken. An estimate of the best reduction of f_{ML} achievable by a local step is given by $\delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}_{\mu=0})$ where $\delta\mathbf{p}_{\mu=0}$ is the unregularized step computed from Eq. (32) with $\mu = 0$. The local objective is then to reduce the maximum likelihood term by:

$$\delta f_{\text{ML}}^{\text{target}} = \varepsilon \delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}_{\mu=0}) \quad (35)$$

where $0 < \varepsilon < 1$ (in practice we take $\varepsilon \simeq 0.4$). To approximately achieve this objective, we search for the value of μ such that $\delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}_\mu) \simeq \delta f_{\text{ML}}^{\text{target}}$. To that end, we use a simple bisection algorithm with automatic updating of the bounds. As a safeguard, the step derived from local minimization of the regularization alone is always computed first:

$$\delta\mathbf{p}_\infty = \lim_{\mu \rightarrow \infty} \delta\mathbf{p}_\mu = -(\mathbf{H}_{\text{prior}} + \eta \mathbf{I})^{-1} \cdot \mathbf{g}_{\text{prior}} \quad (36)$$

then, if $\delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}_\infty) \leq \delta f_{\text{ML}}^{\text{target}}$, the step $\delta\mathbf{p}_\infty$ is taken (the exact value of μ is disregarded).

2. Otherwise, *i.e.* when $f_{\text{ML}}(\mathbf{p}) < f_{\text{final}}$, then μ is chosen so that:

$$f_{\text{ML}}(\mathbf{p}) + \mu f_{\text{prior}}(\mathbf{p}) + \delta f_{\text{ML}}^{\text{quad}}(\delta\mathbf{p}_\mu) + \mu \delta f_{\text{prior}}^{\text{quad}}(\delta\mathbf{p}_\mu) \simeq f_{\text{final}}$$

this ensures that once the parameters become compatible with the data, the regularization is minimized whereas the level of agreement between the model and the data is kept appropriately constant.

3. Finally, whatever the way μ is chosen, we always compute the step $\delta\mathbf{p}_\mu$ according to Eq. (32) and we use the trust-region strategy to update the value of Δ and to decide whether or not the step is acceptable. If not, the trust-region is too large and must be reduced; otherwise, the step is taken and the trust-region radius Δ may be enlarged if the reduction $\delta f_{\text{ML}}^{\text{quad}} + \mu \delta f_{\text{prior}}^{\text{quad}}$ predicted by the quadratic approximation is in good agreement with the actual variation of $f_{\text{ML}} + \mu f_{\text{prior}}$ after the step has been taken.

3.8. Achieving global optimization

The algorithm devised so far may be able to avoid some small local minima (thanks to the prevailing role played by regularization) but is unable to insure global optimization of the highly non-quadratic penalty $f_{\text{ML}} + \mu f_{\text{prior}}$. Expecting that the global minimum yields a good phase estimate for a given data set, one has to consider different initial phases, which are improved by the optimization algorithm until convergence to a local minimum. Among all the local minima found, the deepest one corresponds to the best solution so far. Owing the potentially large number of parameters to describe the phase, about $(D/r_0)^2$, this strategy may however take a very long time to reach the global optimum.

Fortunately, we have been able to make use of the following properties (observed for simulated data) to obtain a good phase estimate in a limited computation time:

- We have seen that the value of f_{ML} which measures the agreement of the model with the data is also a good indicator of the quality of the recovered phase in the sense that if f_{ML} is small enough, then the corresponding phase may not be the best one but is acceptably close to the true phase. This is not unexpected once it is realized that there are much less parameters to describe the phase, about $(D/r_0)^2$, then there are significative measurements (with sufficient signal to noise ratio).

- Our minimization algorithm only performs local optimization but is quite effective: after a limited number of iterations (well before convergence) it is possible to guess whether or not the algorithm will finally converge to a good minimum by comparing the actual value of f_{ML} to some goodness of fit threshold $f_{\text{ML}}^{\text{min}}$.
- We have seen that globally fitting the data for all available wavelengths from the start of the algorithm (when the phase is still badly estimated) yields many more minima whatever is the regularization level and leads to much more unsuccessful minimizations. On the contrary, after having obtained a coarse phase estimate by fitting a single wavelength, the other wavelengths can be accounted for to achieve a much better phase.

Our global optimization strategy is as follows:

1. choose a threshold value $f_{\text{ML}}^{\text{min}}$;
2. select a random initial phase and perform local optimization by fitting a single wavelength data with a phase model described by the first $\sim (D/r_0)^2$ most significant modes of the Karhunen-Loève basis \mathbf{K} ;
3. if $f_{\text{ML}} \leq f_{\text{ML}}^{\text{min}}$ for the phase obtained after a small number of iterations (typically 50), then proceed with next step; otherwise, restart step 2 with a new random initial phase;
4. continue with single wavelength fit until convergence;
5. introduce other wavelengths.

As shown by the results in next section, this approach confirmed our expectation, with the effective retrieval of the correct phase with no degeneracies.

Other specificities of the problem could be used to speed up our global optimization strategy. For instance, the slow temporal evolution of the aberrated phase and a temporal model could provide an initial phase good enough to only require local optimization.

4. SIMULATIONS

We retrieved independent phase screens, with $\frac{D}{r_0}$ between 6 and 10.33 for the longer wavelength λ_0 . We use the potential lever arm of polychromaticity by retrieving derivated phase screens at smaller wavelengths $\lambda_1 = 0.64 \times \lambda_0$, $\lambda_2 = 0.45 \times \lambda_0$ once having selected the good bowl with λ_0 . This enables us to sense wavefronts with $\frac{D}{r_0}$ as high as 26.6.

4.1. Simulated data

The simulated data are 135×135 pixels speckle images. The Kolmogorov phase screens, with 31×31 pupil degrees of freedom at wavelength λ_0 are generated thanks to a Singular Value Decomposition of the piston-removed phase covariance. Photon noise is accounted for by a Poissonian realization of average 10^5 photons, which corresponds to our current speckle experiments. The resulting images are added a gaussian uniform white noise with $\sigma_{\text{CCD}} = 2 \text{ e}^- / \text{pixels}$ derived from our EMCCD.

We simplified our algorithm in approximating \mathbf{W}_{ML} by the identity matrix, though the noise we simulate on data is not uniform. True data may require a more precise weighting, but this simplification enables faster simulations and good results though the weight is not optimal regarding the statistics of the noise.

4.2. Results

We have plotted the standard deviations of the error between the true and the retrieved phases on Fig. 3. Our algorithm always succeeds in retrieving the true phase, which is a major improvement compared to Irwan and Lane phase retrieval simulations⁷ at $\frac{D}{r_0} = 4$. The results with a single wavelength are already very good, with Strehl ratios better than 0.97. Yet the errors with more wavelengths are slightly better and foremost they are less spread out.

With the phase variations becoming more than 2π in a small pupil area for high $\frac{D}{r_0}$, we encountered some unwrapping degeneracies for about the fourth of our retrieved phases at $\frac{D}{r_0} \geq 10.33$, typically over a single small 3×3 area. We did not plot these phases on Fig. 3 because we would not have the problem by restarting the simulations with a proper use of polychromaticity. Yet we deduce from this error of youth that our current global optimization strategy can't efficiently sense $\frac{D}{r_0} \geq 10.33$ from random initial phases and a single wavelength.



Figure 1. Speckle images with $\frac{D}{r_0} = 10.33$ for wavelength λ_0 , 17.77 for wavelength λ_1 and 26.62 for wavelength λ_2

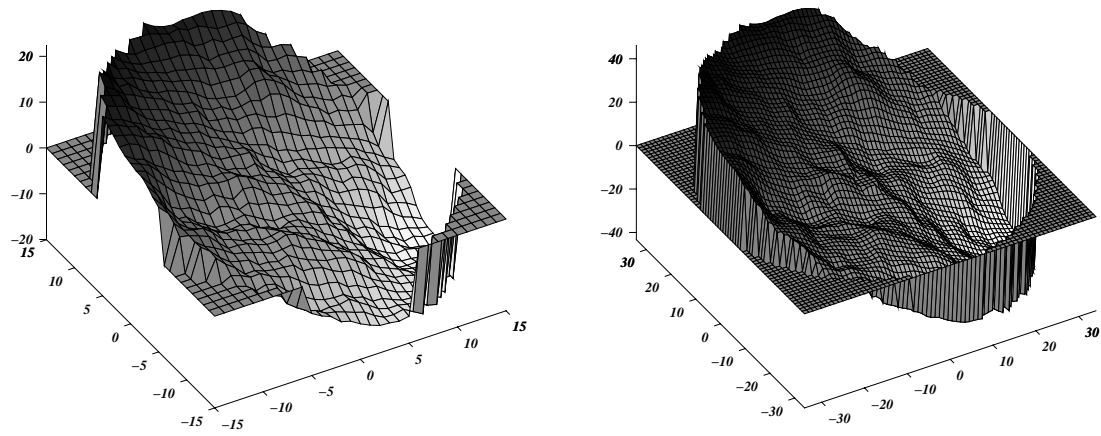


Figure 2. Corresponding Kolmogorov phase screen for λ_0 and the interpolated phase screen for λ_2

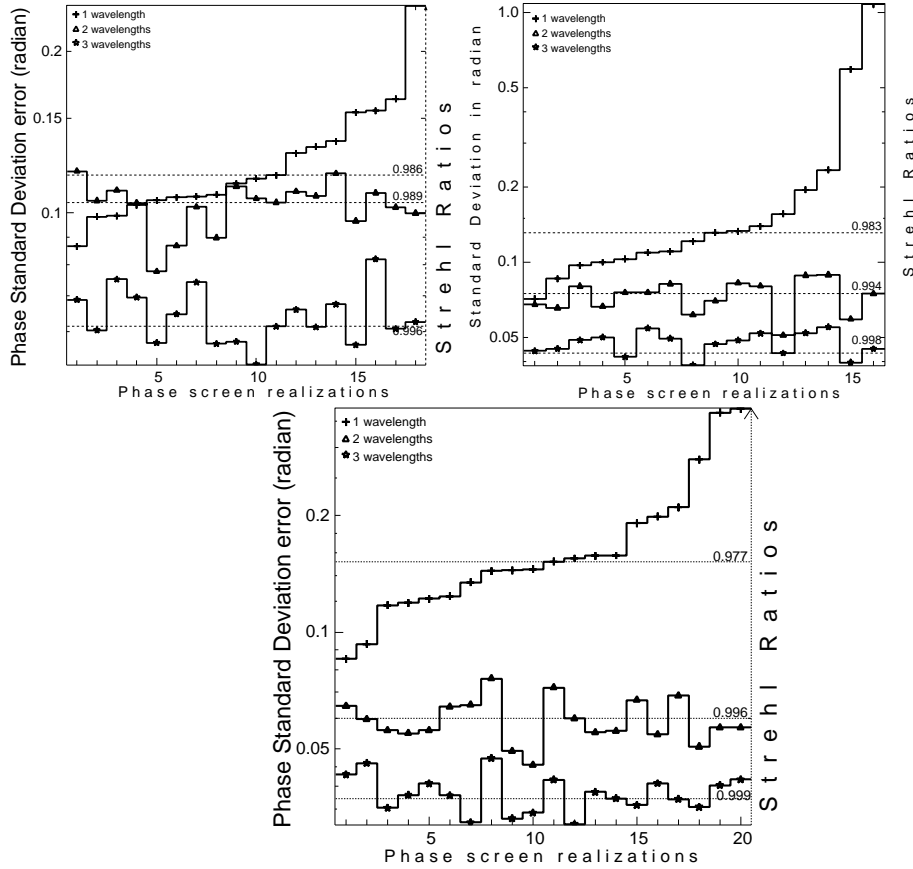


Figure 3. Independent phase std errors for $\frac{D}{r_0} = 6$ (up left), 8 (up right) and 10.33 (bottom) at the longer wavelength λ_0 . The std errors are ordered increasingly with respect to λ_0 . Our global minimization strategy yields the true phase with very small residual errors.

5. APPLICATION TO THE TILT ESTIMATION

In the framework of ELP-OA, we derive the possible improvement of the centroid estimation that phase retrieval enables if we suppose the PLGS is not resolved. It is shown¹ that the atmospheric tip-tilt can be equally recovered from the chromatic differential tip-tilt or centroid. Moreover we use the instinctive assumption that the latter improvement of the centroid error is the same for the tip-tilt so as not to introduce the centroid's expression in the pupil plane.

The Cramer-Rao theorem for a density of probability depending on a parameter θ , gives a lower bound for the variance of the error estimation $E \left\{ \left(\hat{\theta} - \theta \right)^2 \right\}$. For a gaussian spatial distribution of FWHM σ depending on its average position $\tilde{\theta}$, the lower bound on the estimation of $\hat{\theta}$ given N realizations is:

$$E \left\{ \left(\hat{\theta} - \theta \right)^2 \right\} = \frac{\sigma^2}{N}$$

When estimating the centroid (G-tilt) from speckle images, this yields a lower bound for the centroid std error, respect to the FWHM of the images σ_{im} and the number of photons N_{ph} . We now suppose having the true phase for example thanks to phase retrieval. The relative position of all the speckles in the image is known, so that estimating the centroid for instance by means of a maximum of likelihood between our model and the image yields an error on the estimation of the position of a single speckle with N_{ph} photons. Considering a speckle as a gaussian spatial distribution of FWHM σ_{speckle} gives an improvement of $\frac{\sigma_{\text{im}}}{\sigma_{\text{speckle}}}$ for the tilt std error. It is well-known that $\sigma_{\text{speckle}} \propto \frac{1}{D}$ and that for long exposure images $\sigma_{\text{im}} \propto \frac{1}{r_0}$. To obtain the de-tilted short exposure images distribution we should deconvolve the long exposure images

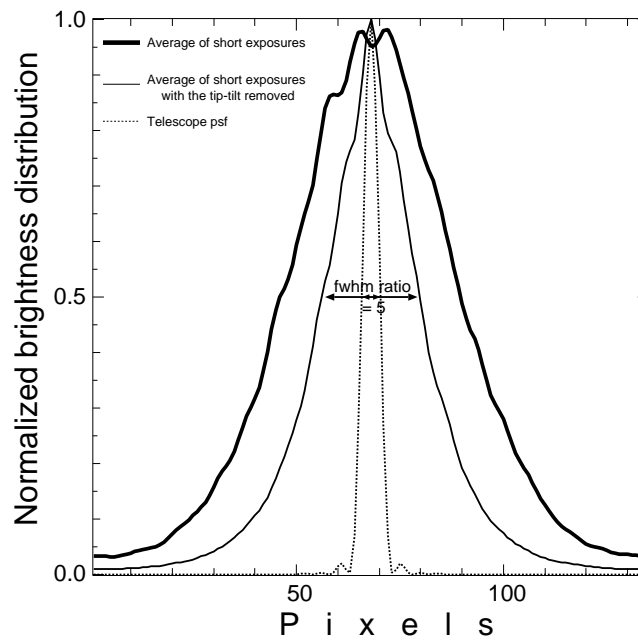


Figure 4. Telescope's psf and average de-tilted short exposure brightness ($\frac{D}{r_0} = 10.33$). The average brightness of a single speckle is considered to correspond to the telescope's psf. The ratio of about 5 between the fwhm is the best achievable improvement of the centroid std error because of the Cramer-Rao lower bound.

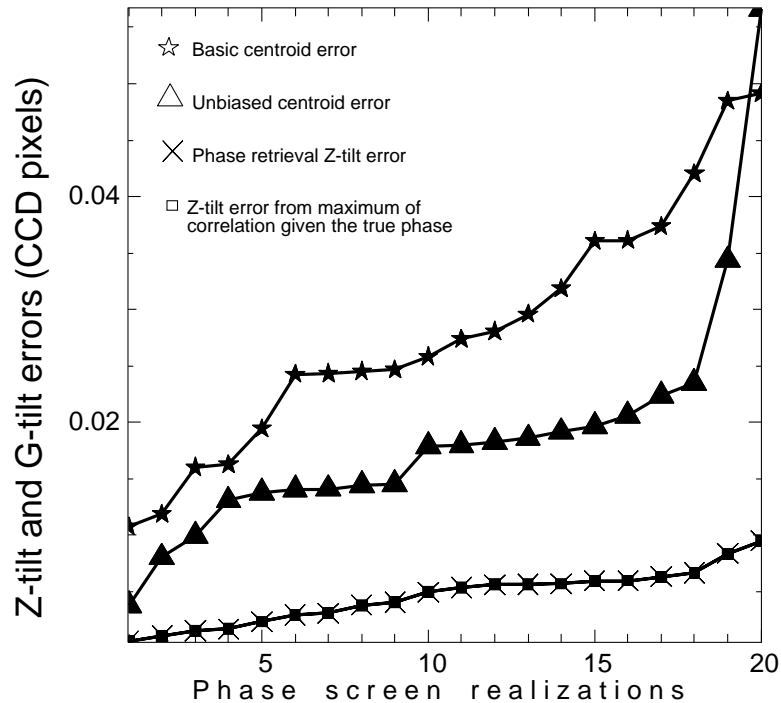


Figure 5. Tip-tilt and centroid errors ($\frac{D}{r_0} = 10.33$ for λ_0): the errors given the true phase and the retrieved phase are almost the same. The improvement of the centroiding std error when the phase is known corresponds to Cramer-Rao statements.

distribution by the centroid distribution. We more simply deduced the FWHM by averaging de-tilted simulated images. For $\frac{D}{r_0} = 10.33$ we get an optimal improvement of 5 on the centroid std error, from the ratio between the deduced and the telescope FWHM. We checked this potential improvement on our simulations.

We plotted the basic centroid, an unbiased centroid given by our best centroiding method, the tip-tilt given by phase retrieval and the tip-tilt from a maximum of correlation given the true phase (on Fig. 5) with increasing order of errors. We found no significant difference between the std error of the latter optimal tip-tilt and of the phase retrieval tip-tilt. Moreover we checked the Cramer-Rao optimal improvement of 5 for $\frac{D}{r_0} = 10.33$ between the two standard deviations of error.

6. CONCLUSION

We have derived our original polychromatic model and our phase retrieval algorithm from a single focal plane. This latter disentangles from the phase retrieval difficulties with both a minimization algorithm with evolving prior constraints accounting for the local roughness and an original global minimization strategy. We have shown from consistent simulations that our algorithm always retrieves the true phase, with residual phase errors corresponding to excellent Strehl ratios even with a single wavelength. We have also pointed out that polychromaticity becomes mandatory for $\frac{D}{r_0} \geq 10$ with our current algorithm. Next steps will focus on various fast ways of initializing the phase, so that only local minimization be required, and we will also retrieve phases from real data. We have finally shown, in the framework of the PLGS, that phase retrieval enables a substantial improvement in the tilt estimation if we suppose the PLGS is not resolved. Future developments have to account for the real brightness distribution of the PLGS.

REFERENCES

1. R. Foy, A. Migus, F. Biraben, G. Grynberg, P. McCullough, and M. Tallon, "The polychromatic artificial sodium star : A new concept for correcting the atmospheric tilt," *Astron. & Astrophys.* **111**, pp. 569–578, 1995.
2. F. Roddier, "The effects of atmospheric turbulence in optical astronomy," in *Progress in optics*, E. Wolf, ed., pp. 281–376, North Holland, Amsterdam, 1981.
3. C. Pichon and E. Thiébaud, "Non-parametric reconstruction of distribution functions from observed galactic discs," *Monthly Notices of the Royal Astronomical Society* **301**, pp. 419–434, 1998.
4. E. Thiébaud and J.-M. Conan, "Strict a priori constraints for maximum-likelihood blind deconvolution," *J. Opt. Soc. Am. A* **12**, pp. 485–492, 1995.
5. S. F. Gull, "Developments in maximum-entropy data analysis," in *Maximum Entropy and Bayesian Methods*, J. Skilling, ed., pp. 53–71, Kluwer Academic, Dordrecht, 1989.
6. R. G. Lane, "Methods for maximum-likelihood deconvolution," *J. Opt. Soc. Am. A* **13**(10), pp. 1992–1998, 1996.
7. R. Irwan and R. Lane, "Phase retrieval with prior information," *J. Opt. Soc. Am. A* **15**, No. 9, pp. 2302–2311, 1998.
8. A. A. Michelson and J. Benoit, "Valeur du mètre en longueurs d'ondes," *Trav. Mem. Bur. Int. Poids Mes* **11**, 1895.
9. E. Thiébaud, "Introduction to image reconstruction and inverse problems," in *Optics in Astrophysics*, R. Foy and F.-C. Foy, eds., NATO ASI, Kluwer Academic, 2005.
10. E. Wallner, "Optimal wave-front correction using slope measurements," *J. Opt. Soc. Am. A* **73**, No. 12, pp. 1771–1776, 1983.
11. R. Lane and M. Tallon, "Wave-front reconstruction using a shack-hartmann sensor," *Applied Optics* **31**, No 32, pp. 6902–6906, 1992.
12. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 2nd ed., 1992.
13. J. J. Moré and D. C. Sorensen, "Computing a trust region step," *SIAM J. Sci. Stat. Comp.* **4**(3), pp. 553–572, 1983.
14. J. Skilling and R. K. Bryan, "Maximum entropy image reconstruction: general algorithm," *Monthly Notices of the Royal Astronomical Society* **211**, pp. 111–124, 1984.

New Generation Photon-Counting cameras: Algol and CPNG

Alain Blazit,¹ Xavier Rondeau,² Éric Thiébaud,² Lyu Abe,³ Jean-Claude Bernengo,⁴ Jean-Louis Chevassut,¹ Jean-Michel Clausse,¹ Jean-Pierre Dubois,² Renaud Foy,^{2,5} Denis Mourard,¹ Fabien Patru,¹ Alain Spang,¹ Isabelle Tallon-Bosc,² Michel Tallon,² Yves Tournéur,⁴ Farrokh Vakili³

¹Observatoire de la Côte d'Azur, GEMINI, Avenue Copernic, 06130 Grasse, France

² Université de Lyon, Lyon, F-69000, France; Université Lyon 1, Villeurbanne, F-69622, France; Centre de Recherche Astronomique de Lyon, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval cedex, F-69561, France; CNRS/UMR-5574; Ecole Normale Supérieure de Lyon, Lyon, France.

³ LUN, Université de Nice, Parc Valrose, 06108 Nice cedex 02, France

⁴ CCQ-UCBL, Université de Lyon, Lyon, F-69000, France; Université Lyon 1, Villeurbanne, F-69622, France and

⁵Observatoire de Haute Provence - CNRS, 04870 Saint-Michel l'Observatoire, France*

Algol and CPNG are new generation photon counting cameras developed for high angular resolution in the visible by means of optical aperture synthesis and speckle interferometry, and for photon noise limited fast imaging of biological targets. They are intensified CCDs. They have been built to benefit from improvements in photonic commercial components sensitivity and personal computer workstations processing power. We present how we achieve optimal performances (sensitivity and spatio-temporal resolution) by the combination of proper optical and electronics design, and real-time elaborated data processing. The number of pixels is 532×516 and 1024^2 read at a frame rate of 262 Hz and 100 Hz for CPNG and Algol respectively. The dark current is very low: $5.5 \times 10^{-4} \text{e}^{-}/\text{pixel/s}$. The saturation flux is ≈ 7 photon events/pixel/s. Quantum efficiencies reach up to 36% and 26% in the visible with the GaAsP photocathodes and in the red with the GaAs ones respectively, thanks to the sensitivity of the photocathodes and to the photon centroiding algorithm; they are likely the highest values reported for ICCDs.

© 2007 Optical Society of America

OCIS codes: (040.0040) Detectors; (030.5260) Photon counting; (100.6640) Superresolution.

1. Introduction: needs and specifications

Photon counting detectors aim at providing images characterized both by particularly low readout and thermal noises, and by fast frame rates. They are dedicated to imaging of low light level sources. In this paper we mean by photon counting detector a device able to detect the position and arrival time of any single photon-event in the image. This is different from a CCD which counts the only integrated number of photons stored per pixel.

We report the specifications and performances of the new generation Algol and CPNG cameras which we have developed. They are intensified CCDs. They are dedicated both to high angular resolution imaging in astrophysics and to fast imaging at low light level in cytobiology.

Since 1974 we have used photon counting cameras to observe astrophysical sources at high angular resolution, mostly

with Michelson interferometry and speckle interferometry. For this purpose, we had built CP40 [1, 2], a previous generation of intensified CCD (ICCD) with 1536×1152 pixels read at 50 Hz per frame. Its quantum efficiency had decreased from $\approx 8\text{-}10\%$ down to $\lesssim 4\%$ because of aging. The resulting limiting magnitude at the GI2T-REGAIN interferometer [3] was $\lesssim 6$. With the SPID speckle interferometer [4], it ranged from $\lesssim 10.5$ to ≈ 16.5 for spectral resolutions of 3000 and 300 respectively. This sensitivity is markedly too low by at least 2 astronomical magnitudes for many high priority science cases of high angular resolution observations, as e.g. extragalactic sources and pre-Main-Sequence stars.

Interferometric observations in the visible need detectors faster than 50Hz in order to get the optimal signal to noise ratio on the interference contrasts under the most common turbulence conditions. Typically a 250 Hz frame rate is required. The spatial resolution must be at least 800×800 pixels (px) for VEGA [5, 6] interferometer to correctly sample the interference fringes and to have the whole image extension in the camera field. For SPID $\sim 1000 \times 1000$ px is required

*Corresponding author: alain.blazit@obs-azur.fr

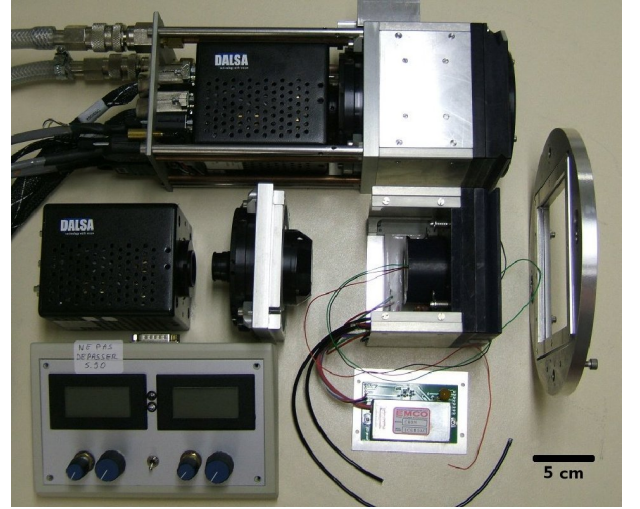
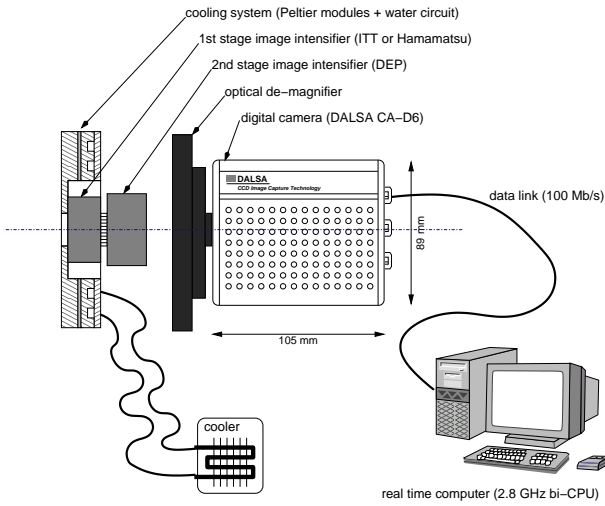


Fig. 1. Left: Hardware design of CPNG. Right: an assembled (top) and a dismantled (bottom) CPNG camera.

for its 4 simultaneous image channels. This can be achieved by means of super resolution processing techniques. In order to observe faint sources, we need a noise level lower than 2500 shots/s. We also need better operability (low weight, automatized cooling).

In life sciences, photon counting imaging is now a routine technique in many fields, such as experimental cancer research. Most studies make use of the large number of insects, bacteria and marine animals expressing luminescent proteins, which are absent in mammals. The injection of luminescent cells into an animal [8], or the expression luminescent proteins in a cell [9] provides a signal detectable with a high specificity. Specific photoproteins can probe calcium ions (aequorin), nicotinamide adenine dinucleotide (bacterial luciferase), superoxide free radicals (lucigenin) or adenosine triphosphate (ATP) (firefly luciferase). Photon counting cameras are routinely used to monitor luminescent tumors in small animals [10, 11]. In our university, we recently developed a photon counting set-up after the pioneering study of Maechler et al. [12]. Cells transfected to express the firefly luciferase emit photons proportionally to the available ATP concentration. The light is collected through a high numerical aperture microscope objective and directed towards the cathode of a photomultiplier. In these experiments at the cellular level, the main information is carried by the temporal signal while the cells are submitted to different experimental conditions [12]. The challenge consists in replacing the photomultiplier by a photon imaging device without losing the real time monitoring signal of photon emission [13]. At high magnification, a microscope has a poor numerical aperture so that it does not require a high spatial resolution; but the time occurrence of photon emission may carry useful information. These requirements may not be fulfilled by a classical integrating camera.

Photon counting cameras made in industry are either too slow (TV speed) or with a relatively low quantum efficiency (using S20 photocathode instead of GaAs or GaAsP). For example, J. L. Gach et al. [14] built a sensitive camera, but not fast enough, only 80 img/s at maximum speed. Electron

Multiplier CCDs (EMCCD), also referred as Low Light Level CCDs (LLLCCD) rely on a new CCD technology [15, 16] which appeared a few years ago. EMCCD's have an amplification register before the readout amplifier which provides intensification of the charge packets by avalanches between electrodes of the register. The gain is sufficient for photon counting operation for astrophysical applications [17] as well as for biological ones [19]. J. L. Gach et al. have already tested such a camera for astronomical applications [20]. We also made some experiments at Observatoire de Haute Provence (OHP) in collaboration with the LISE team, using Andor iXon cameras and at Lyon Observatory with a Roper Scientific *Cascade* camera. Clock induced charges appears in the CCD array, they are amplified like the photoelectrons and produce noise. We have measured that the EMCCD architecture produces $0.5 e^-$ per pixel per frame; this is above our required specification by at least a factor 100. Moreover, with EMCCD's, it is impossible to perform super resolution: detected photon-events spread over only one pixel which prevents better position determination. Finally, existing cameras based on this technique are not fast enough for our needs. Since current EMCCD technology does not match our requirements, we have designed and built Algol and CPNG ICCDs to match the requirements of both domains of biological and astronomical applications.

In the following, we describe the hardware design of the cameras, we explain the real-time data processing and, finally, we present the measured performances of our CPNG cameras.

2. Hardware design

To enable faster build up, flexible data processing and easier upgrade and maintenance, we choose a modular hardware design (cf. Fig. 1) with standard components.

A. Photonic components

In an ICCD, the effective quantum efficiency (QE) and the spectral response are determined by the characteristics of the

first intensifier photocathode. Depending on the readout noise of the CCD, additional secondary intensifiers may be required to provide sufficient light amplification to overcome the CCD noise. Algol and CPNG are designed from commercial on-the-shelf components.

Our principal need is sensitivity and we selected first stage image intensifiers from their photocathode. As a single photocathode cannot completely cover our spectral range, we decided to build cameras with two different first stage intensifiers:

- Hamamatsu V8070U-64 with GaAsP photocathode providing QE between 40 and 50 % in the blue-visible region, 425–650 nm, and a resolution of 51 line pairs per millimeter (lp/mm);
- ITT FS9910 with GaAs photocathode providing QE between 30 and 40% in the red domain, 520–850 nm, and a resolution of 64 lp/mm.

The additional secondary intensifier, which provides us with additional gain, is a DEP XX1450VN, with internal power supply. It has 58 lp/mm, which does not reduce the achievable resolution.

The useful diameter of the photocathodes of the selected light amplifiers is 18 mm.

To focus the intensified image on the CCD, we use a Rodenstock Heligon coupling lens (F/1.55) with adjustable image magnification. For comparable light loss, lens coupling avoids image distortion and micro dislocations by fibre optics coupling. Moreover, the adjustable image magnification is of particular interest for proper sampling of the events. This is required to achieve sub-pixel resolution as explained in section 3.

The readout cameras are coupled to the lens by their C-mounts, and can be exchanged with other cameras as needed. For our applications, we have chosen the following readout cameras (mainly based on their frame rate and number of pixels):

- for SPID and biological applications, the CPNG camera is equipped with a DALSA CA-D6; it is a 8-bit digital CCD camera with 532×516 pixels and a frame rate of 262 img/s;
- for VEGA [5, 6], the Algol camera is equipped with a DALSA 1M150; it is a 8-bit digital CMOS camera with 1024×1024 pixels from which we use only 800×600, and a frame rate of 50 or 100 img/s.

B. Surroundings

To reduce the thermal current noise, the first semi-conductor photocathode must be cooled (noisy events from the photocathode of the second intensifier undergo insufficient amplification so that thermal emission of the second intensifier can be neglected). In our cameras, the cooling is ensured by 10 thermo-electric Peltier modules with cold water circulation for heat dissipation. The temperature is set to -15°C for GaAsP photocathodes and -35°C for GaAs ones (section 4 A). It is

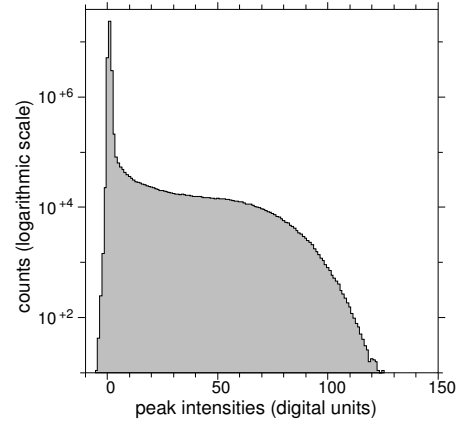


Fig. 2. Histogram of the intensities of the peaks detected in the CPNG readout CCD images.

measured and tuned with a 0.3°C accuracy thanks to a programmable temperature controller. The input window of the first intensifier can be flushed by dry air circulation to remove moist and avoid frost formation. The gains of both intensifiers are manually or computer controlled.

C. Data processing components

Photon counting data processing of the intensified images consists in the detection and localization of the photon-events. With our fast frame rates, it is not possible to save all pixel values. However, only a small fraction of the pixels contains significant signal. Real-time detection of the photon-events is required to reduce the amount of data. Owing to the high pixel rate (up to 72×10^6 pixels/s for CPNG), such a processing is CPU demanding; nevertheless we have developed fast enough algorithms to perform this computation with a 3 GHz Intel Pentium 4 without dedicated programmable electronic devices (such as DSP or FPGA). The computational tasks are:

- images acquisition, without any loss;
- preprocessing of images: de-interleaving and background subtractions;
- photon-events detection and sub-pixel localization;
- real-time photon-events visualization to survey data acquisition and proper camera operation.
- storage of photon events positions and intensities in a 3×3 or 5×5 pixels boxes for posterior processing.

3. Optimal detection and fine centering of the photon-events

The intensified events are seen by the CCD as more or less bright spots. The *energetic distribution* of local maxima in the raw ICCD images is shown in Fig. 2. The peak around 0 DU (Digital Units) in the histogram is mostly due to the readout noise of the CCD. The tail of the distribution at high DU values corresponds to the intensified events.

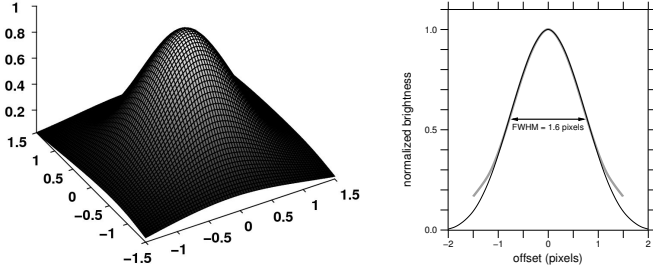


Fig. 3. Photoevent shape. Left: Measured mean 2-D spatial brightness distribution of photon-events on the readout CCD of CPNG. Right: mean 1-D profile of real events (thick gray curve) and B-spline function (thin black curve).

The events *spatial brightness distribution*, due to the intensifier stages and the optical coupling, can be a key point regarding detection and super-resolution (sub-pixel localization of events). On one hand, if an event spreads over less than a CCD pixel, then its detection among the CCD background noise is easier because the peak pixel intensity is maximized, but its position cannot be more precise than a pixel. On another hand, if the event is larger than a CCD pixel, then its spatial spreading makes it possible to achieve sub-pixel resolution by means of centroiding techniques, but the detection efficiency is reduced by the flux dilution with respect to the CCD readout noise. Moreover, wider events yield a larger so-called *photon-counting hole* [2, 21, 22], that is a wider area where only one photon can be detected, which limits the saturation flux of the detector. Since we wanted to achieve reliable sub-pixel resolution whereas keeping good detection efficiency and a good saturation flux, we designed the hardware to have the proper trade-off between these effects.

A. Photon-event shape

As shown by Fig. 3, the measured average profile of the photon-events. It has been obtained by averaging a few ten thousands of recentered (at $1/100^{th}$ pixel precision) events. The full width at half maximum (FWHM) depends on three factors: the spread of the photon burst at the output of the intensifiers, the focalization by the coupling lens, and the CCD pixel size. Assuming that the 2-D brightness distribution $s(\Delta x, \Delta y)$ is separable (which is well verified in practice as can be observed from Fig. 3), it can be written as $s(\Delta \mathbf{x}) = s(\Delta x) s(\Delta y)$, where $\Delta \mathbf{x} = (\Delta x, \Delta y)$ are the coordinates with respect to the peak of the photon-event. Further assuming that it is approximately Gaussian (which can also be observed from Fig. 3) with FWHM ω and that the response of the CCD is uniform within each pixel, then the 1-D distri-

bution writes:

$$\begin{aligned} s(\Delta x) &= \int_{\Delta x - a/2}^{\Delta x + a/2} \exp\left(-\log 16 \frac{u^2}{\omega^2}\right) du \\ &= \frac{1}{2} \left[\operatorname{erfc}\left(\frac{\sqrt{\log 16}}{\omega} \left(\Delta x - \frac{a}{2}\right)\right) \right. \\ &\quad \left. - \operatorname{erfc}\left(\frac{\sqrt{\log 16}}{\omega} \left(\Delta x + \frac{a}{2}\right)\right) \right] \end{aligned} \quad (1)$$

where a is the size of a pixel and $\operatorname{erfc}(z) = 1 - \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ is the complementary error function.

Using Eq. (1) and by numerical simulations, we found that a good trade-off to preserve the efficiency of the detection whereas achieving sub-pixel resolution and keeping a small *photon-counting hole* is to tune the events FWHM in between 1.5 and 2 CCD pixels. Such events roughly spread over 3×3 CCD pixels. With our hardware settings, the actual shape of the real photon-events matches this requirement with $\text{FWHM} \simeq 1.6$ pixels as shown by Fig. 3. After having tuned the focalization of the events with the optical coupling, we have measured that the shape of the events is almost invariant across the whole field of view.

In the following, we will show that an estimate of the photon-event shape is required to maximize the sensitivity of the detection and the accuracy of the event localization. Although we can use the real shape of the events as measured from real data, or the gaussian approximation, we have seen that in practice, the cubic B-spline basis function b yields a suitable approximation with a finite size support as shown by Fig. 3:

$$\begin{aligned} s(\Delta x) &\simeq b(\Delta x/w) \\ \text{with } b(u) &= \begin{cases} 2/3 + (u/2 - 1)u^2 & \text{for } |u| \leq 1 \\ 1/6(2 - u)^3 & \text{for } 1 \leq |u| \leq 2 \\ 0 & \text{for } |u| \geq 2 \end{cases} \end{aligned} \quad (2)$$

where $w = \text{FWHM}_s / \text{FWHM}_b$ is the ratio of the FWHM of the event shape over that of the cubic B-spline, which is:

$$\text{FWHM}_b = \frac{4}{3} + \frac{8}{3} \cos\left[\frac{\pi}{3} + \frac{1}{3} \tan^{-1}(3\sqrt{7})\right] \simeq 1.445. \quad (3)$$

B. Maximum likelihood detection

The first correction applied to the raw CCD images is background subtraction. The background level is quite large on the CCD or CMOS cameras in high sensitivity mode, but it remains stable. So we have measured it by averaging few hundreds of images without intensifiers gain. On the DALSA CA-D6, there is a variable residual background after subtraction of the mean background. It is almost constant along a column. Since under low photon flux, most of the CCD is dark, this columnwise residual background can be easily estimated and fixed. To that end, we subtract to every column (of every image) the level of the peak in the histogram of pixel values along that column.

We use the mean event shape approximated in Eq. (2) to form the convolution kernel.

Owing to the shift-invariant point spread function of our ICCD, it is possible to model the corrected images $d(\mathbf{x})$ as:

$$d(\mathbf{x}) = m(\mathbf{x}) + e(\mathbf{x}) \quad (4)$$

where $\mathbf{x} = (x, y)$ is the position on the CCD, $e(\mathbf{x})$ accounts for noise and modelization errors and $m(\mathbf{x})$ is the *noiseless* model which writes:

$$m(\mathbf{x}) = \sum_k \alpha_k s(\mathbf{x} - \mathbf{x}_k) \quad (5)$$

where \mathbf{x}_k is the central position of the k -th event of the image and α_k is its brightness. Approximating the noise and modelization errors by white noise of variance $\text{Var}[e(\mathbf{x})] = \sigma_{\text{CCD}}^2$, the best model parameters (i.e. brightness and position of the events), in the maximum likelihood sense, would be obtained by minimizing:

$$\chi^2 = \frac{1}{\sigma_{\text{CCD}}^2} \sum_{\mathbf{x}} [d(\mathbf{x}) - m(\mathbf{x})]^2 \quad (6)$$

with respect to the α_k 's and the \mathbf{x}_k 's. A global fit of the image is however too CPU demanding and would not be applicable in real-time. Nevertheless, at low photon counts, the events should be well separated in the observed image. If this is not the case, the overlapping of events would prevent their individual detection, according to the definition of the saturation flux for such a detector (see Section 5 A). Thus instead of minimizing the criterion given by Eq. (6) over the whole image, it is sufficient to apply the maximum likelihood principle *locally* around each event. For the k -th event of the image, this means that one has to minimize:

$$\chi_{\text{local}}^2(\mathbf{x}_k, \alpha_k) = \frac{1}{\sigma_{\text{CCD}}^2} \sum_{\mathbf{x} \in \mathcal{S}_k} [d(\mathbf{x}) - \alpha_k s(\mathbf{x} - \mathbf{x}_k)]^2 \quad (7)$$

where \mathcal{S}_k is the event spatial support. In our case, \mathcal{S}_k is the 3×3 or 5×5 pixel box centered at the event peak value. Since the response of our ICCD is shift-invariant, we have:

$$\mathbf{x} \in \mathcal{S}_k \Leftrightarrow \mathbf{x} - \mathbf{x}_k \in \mathcal{S}$$

where \mathcal{S} is the support of the shape function s and χ_{local}^2 becomes:

$$\chi_{\text{local}}^2(\mathbf{x}_k, \alpha_k) = \frac{1}{\sigma_{\text{CCD}}^2} \sum_{\Delta \mathbf{x} \in \mathcal{S}} [d(\mathbf{x}_k + \Delta \mathbf{x}) - \alpha_k s(\Delta \mathbf{x})]^2. \quad (8)$$

Partial minimization of the criterion in Eq. (8) with respect to α_k only yields the most likely k -th event's brightness given its position:

$$\alpha_k^+(\mathbf{x}_k) = \frac{\sum_{\Delta \mathbf{x} \in \mathcal{S}} d(\mathbf{x}_k + \Delta \mathbf{x}) s(\Delta \mathbf{x})}{\sum_{\Delta \mathbf{x} \in \mathcal{S}} s(\Delta \mathbf{x})^2}. \quad (9)$$

Since $\sum_{\Delta \mathbf{x} \in \mathcal{S}} s(\Delta \mathbf{x})^2$ is constant, the events selected by maximum likelihood are local maxima of the function:

$$d_s(\mathbf{x}) = \sum_{\Delta \mathbf{x} \in \mathcal{S}} d(\mathbf{x} + \Delta \mathbf{x}) s(\Delta \mathbf{x}) \quad (10)$$

which is the discrete correlation of the detected image by the expected shape of the events. Owing to the small size of \mathcal{S} , this correlation can be computed in real-time. Then the local maxima in $d_s(\mathbf{x})$ yield the positions of the events with a pixel precision.

According to our model, the expected value of the peak intensity of the k -th event is: $\langle d(\mathbf{x}_k) \rangle = \alpha_k s(\mathbf{0})$. Therefore, if local maxima are searched for in the raw image $d(\mathbf{x})$, then the detection signal-to-noise ratio is:

$$\text{SNR}_d = \frac{\langle d(\mathbf{x}_k) \rangle}{\sqrt{\text{Var}[d(\mathbf{x}_k)]}} = \frac{\alpha_k}{\sigma_{\text{CCD}}} s(\mathbf{0}). \quad (11)$$

Under our Gaussian white noise approximation, the variance of the CCD intensity's values in the filtered image $d_s(\mathbf{x})$ is:

$$\text{Var}[d_s(\mathbf{x})] = \sigma_{\text{CCD}}^2 \sum_{\Delta \mathbf{x} \in \mathcal{S}} s(\Delta \mathbf{x})^2, \quad (12)$$

and the peak intensity in the filtered image is: $\langle d_s(\mathbf{x}_k) \rangle = \alpha_k \sum_{\Delta \mathbf{x} \in \mathcal{S}} s(\Delta \mathbf{x})^2$; the signal-to-noise ratio for the detection in the filtered image therefore writes:

$$\text{SNR}_{d_s} = \frac{\langle d_s(\mathbf{x}_k) \rangle}{\sqrt{\text{Var}[d_s(\mathbf{x}_k)]}} = \frac{\alpha_k}{\sigma_{\text{CCD}}} \sqrt{\sum_{\Delta \mathbf{x} \in \mathcal{S}} s(\Delta \mathbf{x})^2}. \quad (13)$$

The detection improvement by using the filtered image can be estimated from:

$$\frac{\text{SNR}_{d_s}}{\text{SNR}_d} = \sqrt{\sum_{\Delta \mathbf{x} \in \mathcal{S}} \left(\frac{s(\Delta \mathbf{x})}{s(\mathbf{0})} \right)^2} \simeq 1.5 \quad (14)$$

where the numerical value has been computed for our hardware and our settings. This means that by using the filtered image, the detection of local maxima is significantly more robust. This is of importance to detect the faintest events.

A threshold must be applied to validate a detection, otherwise local peaks due to the CCD noise could be confused with a real event. The threshold level can be derived according to probabilities. Assuming independent uniform Gaussian statistics for the noise and model errors, the probability that a pixel gets a value higher than ϵ is:

$$\Pr\{d(\mathbf{x}) \geq \epsilon\} = \frac{1}{2} \text{erfc}\left(\frac{\epsilon}{\sqrt{2} \sigma_{\text{CCD}}}\right). \quad (15)$$

To have at most one false detection on an average per image, the threshold level must be such that $\Pr\{d(\mathbf{x}) \geq \epsilon\} \leq 1/N_{\text{pix}}$ where N_{pix} is the number of pixels. For the 516×532 DALSA CA-D6 camera with $\sigma_{\text{CCD}} \simeq 0.5$ DU (Digital Units), this yields $\epsilon \geq 4.4 \sigma_{\text{CCD}} \simeq 2.2$ DU. If the detection is performed in the filtered image, the minimum threshold value can be reduced, by a factor $\simeq 1.5$. Nevertheless, since (i) we have observed that the faintest events are mainly due to remanence, and (ii) we want to perform recentering (which requires sufficient discretization levels), we apply a higher threshold level ($\simeq 5$ DU in practice). With such a high threshold, false events are negligible.

To summarize, our real-time detection method is performed as follows: the raw images are convolved by a 3×3 approximation of the shape of the events; in the filtered image, local maxima above a given threshold are selected and the 3×3 or 5×5 surrounding values are saved for further centering. Our detection method is reminiscent to the well-known technique of matching filter which, as we have shown, yields the optimal local detection providing the shape of the events is known and the noise can be approximated by an uniform independent Gaussian statistics.

C. Fine centering

It is well known that the simple photo-center computed from data restricted to a small area \mathcal{S} is biased toward the geometrical center of \mathcal{S} . Various means to achieve better centering in the case of photon-counting data has been proposed by Michel et al. [23]. It is however possible to derive a centering method from our maximum likelihood approach. The criterion in Eq. (7) can be expanded as:

$$\sigma_{\text{CCD}}^2 \chi_{\text{local}}^2(\mathbf{x}_k, \alpha_k) = \sum_{\mathbf{x} \in \mathcal{S}_k} d(\mathbf{x})^2 - 2\alpha_k \sum_{\mathbf{x} \in \mathcal{S}_k} d(\mathbf{x}) s(\mathbf{x} - \mathbf{x}_k) + \alpha_k^2 \sum_{\mathbf{x} \in \mathcal{S}_k} s(\mathbf{x} - \mathbf{x}_k)^2. \quad (16)$$

By replacing α_k by its optimal value $\alpha_k^+(\mathbf{x}_k)$ given by Eq. (9), it comes that the optimal position is obtained by maximizing a criterion which only depends on the position \mathbf{x}_k of the event:

$$\mathbf{x}_k^+ = \arg \max_{\mathbf{x}_k} \frac{\sum_{\mathbf{x} \in \mathcal{S}_k} d(\mathbf{x}) s(\mathbf{x} - \mathbf{x}_k)}{\sqrt{\sum_{\mathbf{x} \in \mathcal{S}_k} s(\mathbf{x} - \mathbf{x}_k)^2}}. \quad (17)$$

To achieve sub-pixel resolution, optimization of criterion in Eq. (17) has to be done for fractional positions of the CCD pixel coordinates. Such an optimization can be done by a Newton method (e.g. the trust region Newton algorithm of Moré and Sorensen[25]) or by tabulating the values of $s(\mathbf{x} - \mathbf{x}_k)$ and $\sqrt{\sum_{\mathbf{x} \in \mathcal{S}_k} s(\mathbf{x} - \mathbf{x}_k)^2}$ for a predefined grid of sub-pixel positions. In any case, the event position $\mathbf{x}_k^+ = (x_k^+, y_k^+)$ given by Eq. (17) is slightly biased toward the center of the CCD pixel. Owing to the stability and the stationarity of the detector response, this bias can be easily fixed by means of histogram equalization of the sub-pixel positions [24].

As shown by Fig. 9, the proposed centering algorithm is effective since it is able to recover the micro-channel plate (MCP) structure (the micro-channel pores are separated by $6 \mu\text{m}$ which is about $1/4^{\text{th}}$ CCD pixel after the optical demagnifier). It however relies on the assumed shape of the events and on the assumption that the CCD response is uniform across every individual pixel.

4. Spurious events

We have seen that proper thresholding drastically reduces the probability of false detections due to the CCD noise. Among the detected events there can be other spurious events due to the amplifying stage defects. These spurious events can be

caused by the dark current of the photocathode, by the remanence of the phosphor screens, and by positive ions at the input of MCP.

A. Dark current

Unlike classical CCDs, the dark current of our camera is not due to the silicium of the CCD, but to the GaAs photocathode of the first intensifier. Indeed any thermal electron emitted by this photocathode is amplified as are the photon-electrons. The gain of the first intensifier (10^4) is much larger than the amplitude of the photon-event dynamics, so the thermal electrons of the second intensifier do not sustain enough gain to be detected. We have measured the dark current over the temperature range between 16 and -35°C , and checked it fits the dark current law of a semiconductor.

The dark current n of the GaAs photocathode follows the Fermi-Dirac statistics:

$$n = N_c \exp\left(\frac{E_F - E_c}{k_B T}\right) \quad (18)$$

where N_c is the equivalent state of density in the conduction band, E_F is the Fermi level of energy, E_c is the conduction level of energy, T is the temperature of the semi-conductor and k_B is the Boltzmann constant. The GaAs semi-conductor can be considered as intrinsic, which means that it has as many electrons as holes, from what follows that:

$$E_F - E_c = \frac{1}{2} \left[k_B T \log\left(\frac{N_v}{N_c}\right) - E_g(0^\circ\text{K}) \right] \quad (19)$$

where E_g is the gap of the semi-conductor and N_v is the equivalent state of density in the valence band. Consequently the variations of our camera with temperature only depend on E_g and N_v .

We have fitted the thermal current measurements as a function of $1/T$ with a logarithmic regression for the lowest measured temperatures, where the gap is less sensitive to temperature variation (Fig. 4). We have derived $E_g \simeq 1.5 \text{ eV}$ which matches the physical value of the gap of an GaAs semi-conductor ($E_g(0\text{K}) = 1.52 \text{ eV}$ for intrinsic GaAs [18]). According to our measurements and to the fitted law, we have chosen a working temperature of -35°C for which we measure 0.45 count per frame which is equivalent to a very good thermal current of $5.5 \times 10^{-4} \text{ e}^-/\text{pixel/s}$ or $\approx 2 \text{ e}^-/\text{pixel/hour}$; the very best CCDs cooled to -110 or -120°C reach the same dark current (in addition they have also a readout noise $\approx 2 \text{ e}^-/\text{pixel}$).

B. Remnant events

The remanence of our camera is due to the temporal response of the phosphor screens which causes the same event to be seen on more than one frame. This effect can be easily evidenced by measuring the spatio-temporal correlations of the detected events. Figure 5 shows correlation profiles as function of the spatial distance $\Delta\mathbf{x}$ and for different temporal separations Δt . Since these profiles are computed from real data with a flat field, the profiles should be mostly flat functions if there is no remanence.

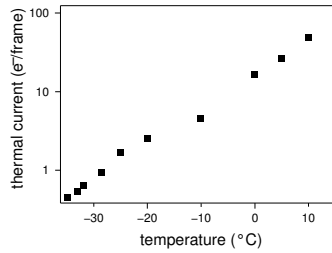


Fig. 4. Thermal current as a function of the temperature of the photocathode of the 1st light amplifier. Squares: measurements.

The excess counts for $|\Delta x| \leq 1$ pixel in the crosscorrelation at $\Delta t = 1$ frame (dashed line in Fig. 5) are due to remnant events which are detected one more time in the next frame (possibly with a slight $|\Delta x| = \pm 1$ pixel offset between consecutive frames due to the CCD noise, to the inhomogeneous microchannel recovery time and to the influence of other close events). The profiles at different time offsets show that remanence is only significant for $\Delta t = 1$ frame and $|\Delta x| \leq 1$ pixel. For this data set, there are about 25 % of remnant events on an average; this is in agreement with the decay time of $\tau \sim 0.7$ ms given by the constructor data sheets for P43 phosphor screens of the light intensifiers.

The depression $\Delta x = \pm 1$ pixel in the autocorrelation ($\Delta t = 0$, solid line in Fig. 5) is due to the inability to detect overlapping events, this is the so-called *photon-counting hole* [2, 21, 22]. Taking into account remanence, the photon-counting hole is a spatio-temporal volume of 3 pixels \times 3 pixels \times 2 frames which matches our expectations from the physical spread and duration of the events and from the detection procedure.

In our data processing, the secondary events in pairs of events inside the photon-counting hole volume are considered as remnant ones and discarded. Note that removal of remnant events does not need to be done in real time and we consider it as a post-processing operation.

C. Ion-events

The last but not least kind of spurious events of our camera are ions in the multiplying stage. Ion-events are created by the detachment of positive ions at the input of MCPs. When we built our cameras, intensifiers with a semi-conductor photocathode were equipped with a special film which performs as an ion barrier to protect the photocathode. Some ions may however cross the barrier and get accelerated back to the photocathode. Photocathodes are coated with aluminium to lengthen their lifetime regarding such ion impacts. The newest intensifiers no more need such films which slightly reduce the quantum efficiency.

Each ion impact creates a burst of electrons, whose amplification in the multiplying stage results in a giant ion-event on the CCD (Fig. 6). Thanks to their large spatial extension, ion-events can be rejected in real time. We have indeed checked that the ratio between the peak intensity of a detected event and the intensities of its four neighbor pixels provides us with a reliable criterion to reject all the ions.

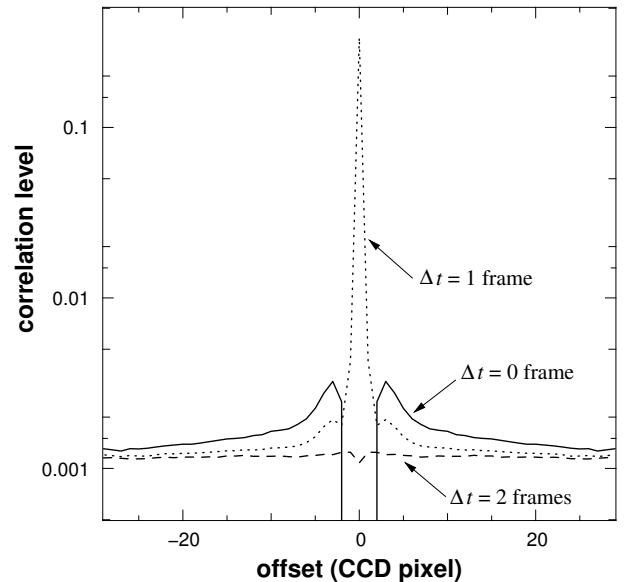


Fig. 5. Spatio-temporal correlation profiles. The curves show the average level of correlation of detected events as a function of the spatial separation and for different temporal separations Δt given in number of 3.8 ms frames.

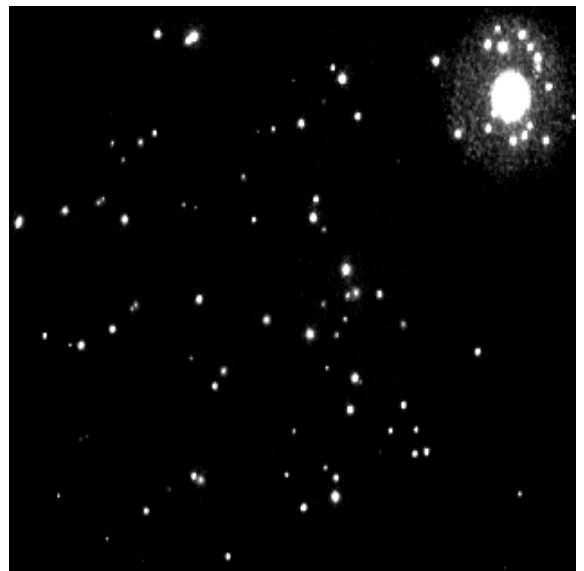


Fig. 6. Single ICCD image showing a large event burst in the top right corner due to an ion.

5. Performances of the camera

A. Photon-counting hole and saturation

The photon-counting hole sets a limitation for the maximum number of photons that can be detected in a single CCD frame. Assuming an uniform incident flux with N incident events on the CCD per frame, it can be shown [22] that the actual number \tilde{N} of detected photons accounting for the loss due to the

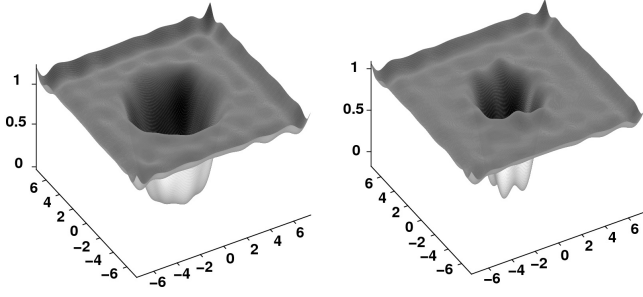


Fig. 7. 2-D mean spatial autocorrelations of photon-events. Left: raw correlation. Right: some possible overlapping of the photon-events is taken into account, which reduces the photon counting hole to its limit 3×3 . Hence the saturation flux (the possible detection area for a single event) is improved.

overlapping of the events reads:

$$\tilde{N} \approx N \exp(-\beta N) \quad (20)$$

where β is the ratio between the photon-counting hole area and the CCD area. On the basis of a maximum of 20 % loss of linearity of the counting mode and with a 3×3 pixels hole (hence $\beta \approx 3.3 \times 10^{-5}$ for CPNG camera), we find that $N \leq 7 \times 10^3$ events/frame, which corresponds to a *saturation* flux of ~ 7 events/pixel/s.

To reduce the effects of overlapping and to limit the size of the photon-counting hole, we have modified our maximum likelihood detection of a single event, by taking into account a non-zero background (namely a constant plus a slope) in the model assumed for the detection. Since s (see Section 3 A) is an even function, this is simply achieved by filtering the raw image by

$$s^*(\Delta\mathbf{x}) = s(\Delta\mathbf{x}) - \langle s(\Delta\mathbf{x}) \rangle_{\Delta\mathbf{x} \in S} \quad (21)$$

instead of s . As shown by the mean autocorrelation in Fig. 7, this simple modification has effectively reduced the photon-counting hole to its smallest size (3×3 pixels) even at relatively high fluxes.

B. Quantum efficiency

To measure the quantum efficiency, we have build low-flux sources at different wavelengths, using temperature stabilized light emitting diodes. Their fluxes have been measured with a calibrated photodiode. We have used a calibrated neutral density filter to reduce the flux so as not to saturate. The global flux precision is about 2 %. Ideally the quantum efficiency should be that of the first photocathode, but some losses are expected.

We have measured the number of real events detected with the different calibration sources by our CPNG cameras (with the GaAsP light intensifier and the GaAs one) as follows. First we have rejected the ion events and also the events due to remanence by means of spatio-temporal crosscorrelations. From the remaining number of detected events, we subtracted the mean counts due to the thermal current. Figure 8 shows

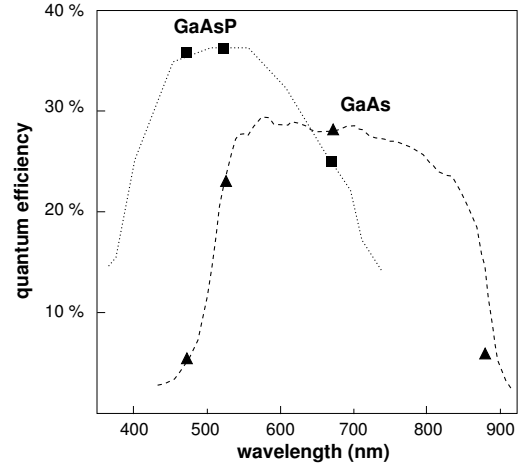


Fig. 8. Quantum efficiency of CPNG/Algol cameras as functions of the wavelength. The squares and triangles indicate our measurements and the curves are the spectral responses of the photocathodes provided by the constructors data sheets and scaled to fit our data. The precision comes from the sources fluxes: 2

the quantum efficiency (QE) measured for our CPNG cameras. It reaches up to 26% for the GaAs photocathode, and 36% for the GaAsP one. This effective quantum efficiency is a dramatic improvement compared to the $QE \sim 4\%$ for the previous generation of photon-counting cameras (CP40).

Such an increase of sensitivity is due to the improved intrinsic sensitivity of the photocathodes, to the reduction of the micro-channels obstruction factor, but also to our careful processing of the raw image data thanks to which very few events are lost by the detection algorithm. The fits in Fig. 8 also show that losses between the photocathode sensitivity and our measured sensitivity are achromatic. The effective quantum efficiency of the photon-counting camera can therefore be expressed as:

$$\eta(\lambda) = \eta_{ph}(\lambda) \eta_{loss}, \quad (22)$$

where η_{ph} is the photocathode sensitivity and η_{loss} is the efficiency of the amplification and detection stages. Taking into account the various independent defects responsible for the sensitivity loss, η_{loss} can be factorized as:

$$\eta_{loss} = \eta_{mc} \eta_{ion} \eta_{sat}, \quad (23)$$

where η_{mc} , η_{ion} and η_{sat} are the attenuation factors respectively due to the microchannels plate obstruction, to the anti-ion film and to the saturation of the counting mode. Since our calibrations have been done under very low photon fluxes, we have $\eta_{sat} \approx 1$. Using our measurements, our model of the saturation loss and the data sheets of the photocathodes, we find that $\eta_{mc} \times \eta_{ion} \approx 0.65$ which agrees with the microchannel plates data sheets.

C. Spatial resolution

To test the ability of our maximum likelihood algorithm to properly estimate event positions with sub-pixel accuracy,

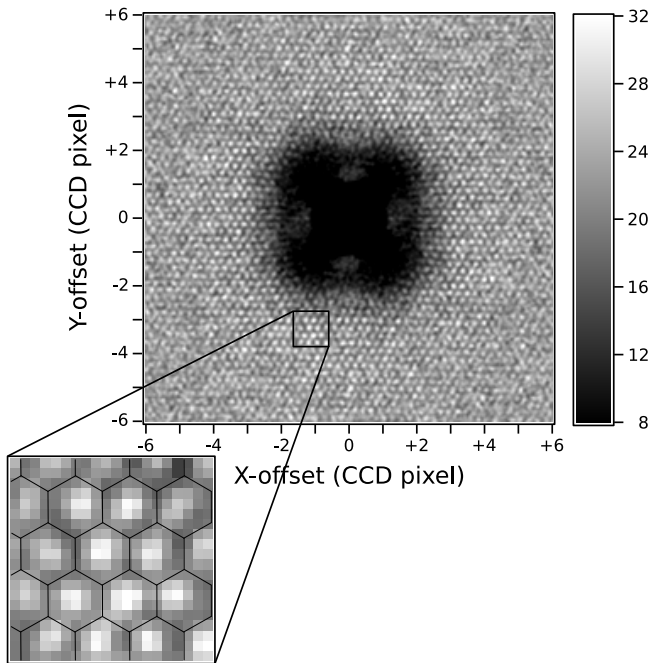


Fig. 9. Mean autocorrelation of detected event positions under flat field and computed with a resolution of a 20th of a CCD pixel. Axis are labeled in pixel units. The accurate event positions were determined by a maximum likelihood fit followed by histogram equalization of the sub-pixel positions to avoid the bias of the centering method. The honeycomb structure of the microchannel arrays (7 fibers per pack) is clearly seen in the 1 squared CCD pixel zoomed area. The dark part around coordinates (0, 0) is due to the photon-counting hole.

we averaged the autocorrelations of photon-counting images under uniform illumination (flat field) and with a super-resolution of $1/20^{\text{th}}$ pixel. The honeycomb structure of the microchannel pores can be clearly seen in Fig. 9 which shows the central part of the average autocorrelation. Taking into account the optical magnification, the $\sim 1/4^{\text{th}}$ CCD pixel size of the honeycomb structure in Fig. 9 matches the separation of $6\mu\text{m}$ of the microchannel pores. In practice, there is no means to determine the position of incident events with a better resolution than that of the microchannels which is much coarser than the accuracy achievable by the fine centering of events. The effective spatial resolution of our photon-counting cameras is therefore limited by the microchannels structure to about $\sim 1/4^{\text{th}}$ CCD pixel. In photon-counting mode, CPNG and Algol are equivalent to 2-D cameras with $\sim 2000 \times 2000$ to $\sim 3000 \times 3000$ effective pixels depending on the optical setup.

6. Conclusion

Today, 6 Algol-CPNG ICCDs have been built. New microchannels and image processing enable us to reach up to 36% and 26% quantum efficiencies in the blue (GaAsP) and in the red (GaAs) respectively; this is 65% of the quantum efficiency of these photocathodes. It is a marked increase in

the sensitivity compared to the previous (CP40) ones. Thanks to the new photon centroiding algorithm which we have developed, there are no traces of the CCD pixels in the super resolution images (1 logical pixel = $1/(4 \times 4)$ CCD pixel), so that the resolution is only limited by the size of the individual microchannels.

These new ICCDs are the most suitable ones for high angular resolution observations in the visible at 8-10 m class telescope and at large interferometers. The increase of limiting magnitudes of our instruments will open new classes of objects to observations at high angular resolution. From mid 2003 to 2005, both Algol cameras were used on GI2T-REGAIN interferometer. Now the spectrograph of REGAIN and the Algol cameras are installed at the CHARA interferometer. SPID is being modified to match the photocathode diameter smaller than the CP40 one.

For biological applications, the high quantum efficiency of the camera will enable us to get images at the optimum resolution given the selected objective, and a frame by frame photon count will provide us with a unique possibility to simultaneously monitor the time event of photon emission, possibly opening a field of investigations on the kinetics of photo-biochemical reactions.

This work was supported by Region Rhône-Alpes (contrat #99.822.195) and by CNRS/INSU. The algorithms and the simulations presented in this article have been implemented in Yorick, a free data processing language written by D. Munro (<http://sourceforge.net/projects/yorick/>).

References

1. A. Blazit, "A 40-mm photon counting camera" Proc. SPIE **702**, 259 (1987).
2. R. Foy, "The photon counting camera CP40", in *Instrumentation for Ground-Based Optical Astronomy, Present and Future* Lick Observatory, eds. (Springer-Verlag, New York, 1988), pp. 589.
3. D. Mourard et al., "Status report on the GI2T interferometer", Proc. SPIE **4838**, 9 (2003).
4. M. Tallon et al., "SPID, a high spectral resolution diffraction limited camera" Proc. SPIE **4007**, 962 (2000).
5. D. Mourard et al., "VEGA: a visible spectrograph and polarimeter for CHARA" Proc. SPIE **6268**, 62683Q (2006)
6. P. Stee et al., "VEGA: a visible spectrograph and polarimeter for CHARA - science cases description" Proc. SPIE **6268**, 62683R
7. R. Foy, D. Bonneau and A. Blazit, "The multiple QSO PG1115 + 08 - A fifth component?", *Astronomy and Astrophysics* **149**, L13-L16 (1985).
8. R.S. Negrin, C.H. Contag, "In vivo imaging using bioluminescence: a tool for probing graft-versus-host disease", *Nat. Rev. Immunol.* **6**, pp. 484-490 (2006).
9. A. Roda, P. Pasini, M. Mirasoli, E. Michelini, and M. Guardigli, "Biotechnological applications of bioluminescence and chemiluminescence", *Trends Biotechnol* **22**, pp. 295-303 (2004).

10. L.F. Greer 3rd, and A.A. Szalay, "Imaging of light emission from the expression of luciferases in living cells and organisms: a review", *Luminescence* **17**, pp. 43-74 (2002).
11. R.T. Sadikot and T.S. Blackwell, "Bioluminescence imaging", *Proc. Am. Thorac. Soc.* **2**, pp. 537-40, 511-2 (2005).
12. P. Maechler, H. Wang and C.B. Wollheim, "Continuous monitoring of ATP levels in living insulin secreting cells expressing cytosolic firefly luciferase", *FEBS Lett.* **422**, pp. 328-32 (1998).
13. H.J. Kennedy, A.E. Pouli, E.K. Ainscow, L.S. Jouaville, R. Rizzuto, G.A. Rutter, "Glucose generates sub-plasma membrane ATP microdomains in single islet beta-cells. Potential role for strategically located mitochondria", *J. Biol. Chem.* **274**, pp. 13281-13291 (1999).
14. J.-L. Gach et al., "Fabry-Pérot Observations Using a New GaAs Photon-counting System", in *Scientific Detectors for Astronomy, The Beginning of a New Era* P. Amico et al., eds. (Kluwer Academic, 2002), pp. 335-339.
15. P. Jerram et al, "The LLLCCD: Low light imaging without the need for an intensifier", in *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications II* **4306**, pp. 178-186, SPIE, 2001.
16. M.S. Robbins and B.J. Hadwen, "The noise performance of electron multiplying charge-coupled devices", *IEEE Trans Electron Devices*, **50**, 1227-1232 (2003).
17. A.G. Basden, C.A. Haniff, and C.D. MacKay "Photon counting strategies with low-light level CCDs", *Mon. Not. R. Astron. Soc.*, **345**, 985-991, (2003).
18. O. Madelung, M. Schulz and H. Weiss, "Intrinsic Properties of Group IV Elements and III-V, II-VI, and I-VII compounds", *Landolt-Bornstein, New Series, Group III*, **22**, Pt. a, Springer, Berlin (1987)
19. C.G. Coates, D.J. Denvir, N.G. McHale, K.D. Thornbury, and M.A. Hollywood, "Optimizing low-light microscopy with back-illuminated electron multiplying charge-coupled device: Enhanced sensitivity, speed, and resolution" *J. Biomed. Opt.* **9**, 1244-1252 (2004).
20. J.-L. Gach, C. Guillaume, O. Boissin and C. Cavadore: "First results of an L3CCD in photon counting mode", *Proc. Astrophysics and space science library* **300**, 611 (2004).
21. E. Thiébaud, "Speckle interferometry with a photon-counting detector", *Astronomy and Astrophysics* **284**, 340-348 (1994).
22. E. Thiébaud, "Avoiding the photon-counting hole in speckle imaging by means of cross-correlation techniques", *J. Opt. Soc. Am. A* **14**, 122-130 (1997).
23. R. Michel, J. Fordham, and H. Kawakami, "Fixed pattern noise in high-resolution, ccd readout photon-counting detectors", *Mon. Not. R. Astron. Soc.* **292**, 611-620 (Dec 1997).
24. M.K. Carter, R. Cutler, B.E. Patchett, P.D. Read, N.R. Waltham & I.G. van Breda, "Transputer-based image photon-counting detector", in *Instrumentation in Astronomy VII*, Proceedings of SPIE **1235**, pp. 644-656 (1990).
25. J.J. Moré and D.C. Sorensen, "Computing a trust region step" *SIAM J. Sci. Stat. Comp.* **4**, pp. 553-572 (1983).

**Imagerie à travers la turbulence :
mesure inverse du front d'onde et centrage optimal**

L'imagerie à la limite de diffraction des télescopes de grand diamètre constitue un enjeu majeur pour l'astrophysique moderne. Mais il faut compenser les effets délétères de la turbulence sur la phase de l'onde électromagnétique.

Je développe un algorithme d'optimisation globale pour le problème inverse non-linéaire de restauration de phase à partir d'une image tavelée, et je montre le bénéfice de la diversité polychromatique (images à différentes longueurs d'onde) pour gérer plus de degrés de liberté. Pour l'Etoile Laser Polychromatique, je montre qu'estimer conjointement toute la phase améliore considérablement l'estimation du déplacement par rapport au barycentre. Je développe aussi une méthode d'interspectre bichromatique pour estimer le déplacement lorsque l'étoile est résolue. Je présente enfin une approche inverse pour traiter en temps-réel les données d'un détecteur à comptage de photons (gain 4 en résolution spatiale).

Mots-clés : problèmes inverses, restauration de phase, centrage, étoile laser

**Imaging through the turbulence :
phase retrieval and optimal centering**

Diffraction-limited imaging with Extremely Large ground-based Telescopes is a major challenge of modern astrophysics. But it requires the correction of the wave-front corrugations due to the atmospheric turbulence.

I present a global optimization algorithm to solve the non-linear inverse problem of phase retrieval from a speckle image, and I show the benefit of polychromatic diversity (images at different wavelengths) to deal with more degrees of freedom. In the Polychromatic Laser Guide Star framework, I address the problem of assessing the chromatic differential tip-tilt with more accuracy than centroids. The benefit of the full phase joint estimation is evidenced, in the case the PLG Star is unresolved. Otherwise I develop an estimator of the star's differential movement, based on the bichromatic cross-spectrum. Finally, I present an optimal real-time data processing of a photon counting detector with an inverse problem formulation (gain of 4 in spatial resolution).

Keywords : inverse problems, phase retrieval, centering, laser guide star
